Rare Vocabulary Production in School-age Narrators from Low-income Communities

JAMIE MAHURIN-SMITH[1], MONIQUE MILLS[2], and RONG CHANG[3]

1. Department of Communication Sciences and Disorders, Illinois State University, Normal Illinois

2. Department of Communication Sciences and Disorders, University of Houston, Houston Texas

3. Khoury College of Computer Sciences, Northeastern University, San Francisco, CA

Corresponding Author:

Jamie Mahurin-Smith, Department of Communication Sciences and Disorders, Campus Box 4720, Illinois State University, Normal IL 61790. E-mail: j.m.smith@ilstu.edu. Phone: 309-438-5308.

Conflict of Interest Statement: The authors have no conflicts of interest to report.

Funding Statement: This study was funded in part by a Social and Behavioral Sciences Small Grant awarded to the second author.

Abstract

Purpose: This study was designed to assess the utility of a tool for automated analysis of rare vocabulary use in the spoken narratives of a group of school-age children from low-income communities.

Method: We evaluated personal and fictional narratives from 76 school-age children from low-income communities (mean age = 9;3). We analyzed children's use of rare vocabulary in their narratives, with the goal of evaluating relationships among rare vocabulary use, performance on standardized language tests, language sample measures, sex, and use of African American English (AAE).

Results: Use of rare vocabulary in school-age children is robustly correlated with established language sample measures. Male sex was also significantly associated with more frequent rare vocabulary use. There was no association between rare vocabulary use and use of AAE.

Discussion: Evaluation of rare vocabulary use in school-age children may be a culturally fair assessment strategy that aligns well with existing language sample measures.

Rare Vocabulary Production in School-Age Narrators from Low-income Communities

Speech-language pathologists need a variety of culturally fair assessment instruments in order to evaluate culturally diverse clients. In this paper, we propose that assessment of children's use of rare vocabulary may be such an instrument. Vocabulary selection is a key element of narrative construction. Analysis of the vocabulary deployed by young narrators – particularly their use of uncommon words – has the potential to serve as a useful strategy for evaluating language skills. The purpose of this paper is to assess the utility of a tool for automated analysis of rare vocabulary use in the spoken narratives of a group of school-age children from low-income communities, with the goals of providing more culturally fair evaluations within that population. In this paper we investigate (1) the tool's alignment with more conventional language assessment strategies, and (2) its capacity to measure vocabulary diversity within the sample as a function of (a) sex and (b) dialect density.

In the sections that follow, we review the assessment challenges faced by children in low-income communities, the importance of narrative in providing individualized language evaluations, the specific importance of vocabulary within narrative, and the potential contribution of rare vocabulary assessment.

Assessment Concerns in Low-income Communities

Achievement gaps are long-standing across the social-class spectrum in United States. Children from low-income communities are chronically underrepresented in gifted education (Hamilton, McCoach, Tutwiler, Siegle, Gubbins, Callahan, Brodersen, & Mun, 2018) and overrepresented in special education in the U.S. (Connor & Boskin, 2001). Scholars observe similar persistent academic inequities between African Americans and other groups because of the confounding influence of race on economic, health, and social well-being in the U.S. African

American children continue to lag behind on some educational measures in the U.S. compared to their Asian American and White peers (Leu, Forzani, Rhoads, Maykel, Kennedy, & Timbrell, 2014; National Center for Educational Statistics [NCES], 2011). In particular, African American boys from low-income backgrounds make up a disproportionately large fraction of K-12 special education caseloads (Sullivan & Bal, 2013), a problem that suggests an urgent need for culturally fair assessment tools.

Although the factors that contribute to these educational disparities are complex and myriad, language skills form the bedrock of academic success. As such, language represents a pressing variable to assess in African American children because they may enter school speaking African American English (AAE)—a variety of language that has been assigned low prestige (Wolfram & Schilling-Estes, 2006). AAE production, though linguistically legitimate and rule-governed, is considered an academic risk factor for multiple reasons. Many studies show that AAE shares an inverse relationship with performance on tests of reading in Standard American English (SAE; Gatlin & Wanzek, 2015; Terry, Connor, Johnson, Stuckey, & Tani, 2016; Craig, Kolenic, & Hensel, 2014). Specifically, students with higher oral and/or written scores on a dialect density measure (DDM), designed to evaluate how frequently children use AAE, may perform worse on measures of reading ability (Craig, Zhang, Hensel, & Quinn, 2009; Craig, Thompson, Washington, & Potter, 2004). This is particularly true for those who struggle with code-switching to MAE. Moreover, teacher-student interactions may be influenced by implicit bias or negative language ideologies toward AAE (Blake & Cutler, 2003).

Accurate assessment of AAE speakers is also a known diagnostic challenge (Craig & Washington, 2000; Hendricks & Adolf, 2017). Across grade levels, differences between home and school language can impede accurate assessment of students' knowledge and skills via

classroom grading practices and standardized tests (Wheeler, Cartwright, & Swords, 2012; Wheeler, 2019). A substantial body of research indicates that failure to accept and accommodate AAE fuels educational disparities and, ultimately, disparities in achievement between MAE speakers and AAE speakers (e.g., Carter, 2010; Hallett, 2015; Stockman, 2010). Effective responses to these disparities require a thorough understanding of formal and informal evaluation strategies for the language of school-age Black children.

In their 2009 paper, Dudley-Marling and Lucas called on educators to reject a "deficit stance," in which children from poor families are assumed to be suffering from cultural and linguistic deficiencies, in favor of acknowledging individual differences and working to identify the unique cultural and linguistic resources available to students regardless of their backgrounds. We thus suggest that a more individualized approach to language assessment among children from low-income communities has the potential to provide more useful results, and that analysis of narratives on topics of interest to individual students may prove particularly valuable for this purpose.

Narrative Performance in Children from Low-income Communities

Studies of narrative development in children from low-income communities have demonstrated that both monolingual and bilingual children improve the quantity and quality of their narrative language over time (Benson, 1997; Price, Roberts, & Jackson, 2006; Bitetti & Hammer, 2016). For example, in their study of 65 African American children, Price and colleagues (2006) found improvement on almost all measures of narrative coherence from pre-kindergarten to kindergarten. Neither parental education nor family income accounted for the variability in this growth.

Collectively, studies of narrative development in school-age children from low-income communities suggest that average sentence length, vocabulary diversity, and narrative coherence improve with age (Mills, 2015a; Mills, 2015b; Mills, 2016; Mills et al., 2017). Although observers expect this kind of upward trajectory in all populations of school-age children, low-income children's language tends to remain less sophisticated than that of mid- and upper-income children (Hoff, 2006). For this reason, group comparisons that combine participants from low-income communities and high-resource communities may obscure observable linguistic strengths among individual low-income children, particularly in view of the widespread "deficit stance" described by Dudley-Marling and Lucas (2009). Therefore, we argue that an explicit focus on variability among narrators from low-income communities has the potential to provide a more holistic perspective.

The Importance of Vocabulary Assessment

Prior research into children's narratives has highlighted both its macrostructural and microstructural elements (e.g., mean length of utterance (MLU), number of different words (NDW), and number of total words (NTW)). In this paper, we consider the specific insight into language abilities afforded by analyzing vocabulary selection. Researchers have reported strong positive associations between children's narrative abilities and their vocabulary skills (Uccelli & Paez, 2007). Vocabulary skills are particularly critical for children because they are malleable abilities reflecting mental representations of a growing array of concepts; children's use of vocabulary within their narratives, then, may provide a window into a domain that is critical for effective language comprehension and use as well as for academic success (McGregor, Newman, Reilly, & Capone, 2002). In this paper, we consider specifically the insight into language abilities afforded by vocabulary selection.

Assessment of children's expressive vocabulary in individualized tasks may be a tactic that is particularly well-aligned with the exhortation of Dudley-Marling & Lucas (2009) to evaluate children with an eye to their individual strengths. Children absorb new vocabulary more easily when it represents topics of interest to them (Mani & Ackermann, 2018) and when the new words are elements of a semantically dense network rather than a semantically sparse network (Borovsky, Ellis, Evans, & Elman, 2016).

Individual differences such as gender also create a challenge in providing culturally fair evaluations of children's vocabulary skills. It is well-documented that girls acquire vocabulary more rapidly during the early years of language acquisition (Huttenlocher, Haight, Bryk, Saltzer, & Lyons, 1991), and some research has suggested that this advantage persists as girls reach preschool (Craig & Washington, 2004) and school age (Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004). An argument might be made that having established a beachhead, so to speak, with their early vocabulary advantage, girls would be well-positioned to build on those existing skills as they tackle the task of acquiring more abstract, literate, and uncommon vocabulary. Some evidence exists in support of this perspective; Kaushanskaya, Gross, and Buac (2013) reported that in their sample of school-age children, the girls demonstrated stronger retention of novel vocabulary.

One might, however, offer the counterargument that boys' vocabulary catches up to girls' as they mature. Within the literature on speakers who are African American and/or members of low-income communities, the evidence is equivocal. For example, Barbu and colleagues (2015) indicated that effects related to sex may also vary by socioeconomic status (SES), with a more pronounced advantage for lower-SES girls in comparison with lower-SES boys. In contrast, Allison and colleagues (2011) found significantly stronger receptive vocabulary skills among the

boys in their sample of 30 low-SES African-American preschoolers. To our knowledge, no empirical evidence exists on sex-related vocabulary differences among school-age children from low-income communities. Given the overrepresentation of African American boys on special education caseloads (Connor & Boskin, 2001), we argue that understanding the range of vocabulary skills among students in general education classrooms is an important step toward accurate identification of children with special needs.

The vast majority of the available evidence on vocabulary acquisition in school-age children relies on standardized vocabulary assessments; these tests often emphasize less common words. Assessing children's use of rare vocabulary directly, based on their individual narratives, offers a promising, less traditional approach

Rare Vocabulary as an Assessment Strategy

Mahurin-Smith and colleagues (2013) briefly described their use of an automated analysis strategy for assessing children's use of rare vocabulary in conversations and narratives; they tallied instances of words that appeared ≤15 times in a corpus of 1.2 million words spoken by a sample of 876 school-age children aged 5-12. These findings were extended in a 2015 paper by Mahurin-Smith, DeThorne, and Petrill, in which the authors reported that results from their tool for assessing rare vocabulary use, Wordlist for Expressive Rare Vocabulary Evaluation (WERVE) aligned with both standardized test results and existing language sample vocabulary measures (e.g., number of different words (NDW), number of total words (NTW)). Both of these investigations described results from subsets of the Western Reserve Reading and Math Project, a sample of predominantly White children with educated parents. More specifically, the 2013 paper described the use of WERVE in a cohort of 114 children, of whom 104 were white, 8 were African American, and 2 identified themselves as "other." Parental education was used as a

proxy for SES in this study. Every parent in the sample had at least a high school diploma; furthermore, 81% of them had completed at least some college and 28% of them had completed graduate degrees.

In contrast to existing measures of vocabulary rarity (notably the three-tiered classification system for vocabulary (McKeown & Beck, 2004), which is widely used to classify words used in written texts produced for children), WERVE is based entirely on spoken language produced by school-age children. As illustrated in the Appendix, most of the words it identifies as rare in children's spontaneous spoken language samples would probably be classified as Tier 2 vocabulary in written texts for younger school-age children, though WERVE includes instances of Tier 3 vocabulary as well.

Following these initial investigations of rare vocabulary in samples of predominantly white middle-class children, a 2017 paper by Mills, Mahurin-Smith, and Steele explored the utility of WERVE in evaluating stories produced by African American children from gifted and general education classrooms with a different demographic profile. Of the children in this 2017 study, 69.8% came from low-SES backgrounds. In this sample, children's use of rare vocabulary predicted gifted status: gifted African American children were more likely to produce rare vocabulary in their personal narratives than were typically developing (TD) African American children. There was no association between rare vocabulary use and children's use of African American English, suggesting that WERVE is a culturally fair tool. These findings raise questions about the potential uses of WERVE in evaluating narratives from a sample of African American children from low-income communities, where the need for culturally fair assessment tools is particularly pressing. Specifically, would WERVE yield useful supplemental information in a sample of general education students, with a higher proportion of participants from low-SES

backgrounds? Would it serve as a culturally fair tool that aligns with established language sample measures? Would it highlight individual differences? Prior work indicated that WERVE could be useful in identifying giftedness, but would WERVE provide meaningful information about language performance in a sample of middle-ability students?

The Current Study

In the preceding sections we have discussed the need for individualized assessment of children from low-income communities, as well as the importance of narrative as an expression of their language skills and the value of vocabulary – specifically rare vocabulary – in evaluating their narratives. For this reason, we opted to investigate personal and fictional narratives constructed by school-age children, using a tool that would allow us to compare the relative rarity of their individual vocabulary selections to a sizable corpus of language produced by other school-age children. To assess the potential utility of rare vocabulary assessment among children from low-income communities, we wanted to determine whether WERVE results would provide a culturally fair supplemental assessment measure that aligned with established language sample measures. In order to address the issue of overidentification among African American males we also elected to look at sex differences in related rare vocabulary use. The present study, then, addresses the following research questions: (1) Does rare vocabulary use align with performance on standardized measures and/or language sample measures? (2) Does rare vocabulary use vary by sex among children in this sample? and (3) What is the relationship between rare vocabulary and AAE production? Based on the findings reported in Mills et al. (2017), we hypothesized that rare vocabulary use would align with performance on standardized language tests and language sample measures. The existing literature suggested that girls might use more rare vocabulary

than boys. We further hypothesized that there would be a null relationship between rare vocabulary and dialect variation, again based on results from Mills et al. (2017).

Method

Participants

The current study describes a supplemental analysis of data originally reported in Mills and Fox (2016), with the addition of new data from a further 26 participants recruited after its publication. This was a convenience sample of children recruited through public elementary schools in central Ohio, obtained to investigate relationships among language, literacy, and social capital in children from low-income communities. We recruited 92 participants into the larger study, including the 66 described in Mills & Fox (2016). For inclusion in the present study, participants needed to be in grades 2-5 and to complete all required assessment tasks. We had information on special education eligibility from 67% of parents, indicating that 16 children were receiving support services for the following: attention deficit disorder (n = 1); English as a Second Language (n = 2); gifted (n = 7); reading (n = 2); speech (n = 2); and unspecified (n = 2). For the current study we excluded these 16 children from consideration in order to focus specifically on typical development. We were left with a sample of 76 children (40 females, 36 males). This study received Institutional Review Board approval from the Ohio State University.

Children attended public elementary schools in central Ohio and were educated in self-contained general education classrooms. A one-page demographic survey was completed by one of the child's parents to gather information about the child's age, race/ethnicity, family SES, and special education placement. On average, children were 9;3 (years;months) and ranged from 7;1 to 11;5. The sample included children from diverse ethnic backgrounds: 88.2% African

American (*n* = 67); 5% African immigrant (*n* = 4); 5% European American (*n* = 4); and 1% Asian (*n* = 1).

Forty-three percent (*n* = 33) of parents provided information about their educational level, which was used a proxy variable to estimate family SES, as follows: 9.2% (*n* = 7) reported having less than 12 years; 18.4% (*n* = 14) reported being high school graduates; 10.5% (*n* = 8) reported having greater than 12 years, but were not college graduates; and 5.3% (*n* = 4) reported being college graduates. For the remaining 56.6% of the sample (*n* = 43), we determined family SES based on school free lunch status. In the two schools that provided our sample, free lunch status was 81.75% and 100%, respectively. All parent reports indicated that children in the sample had never received special education services.

Diagnostic Testing

Two graduate clinicians and the first author administered tests to characterize the overall cognitive and vocabulary abilities of each child. The first author, a certified speech-language pathologist, elicited narratives from each child through language samples and a norm-referenced test of narration. Prior to gathering data independently, graduate clinicians in communication sciences and disorders administered tests to each other and observed the first author assessing research participants. Graduate clinicians then gathered data under close supervision from the first author until they followed assessment procedures accurately and consistently.

All testing occurred in two semi-private rooms in a local elementary school. The Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007) assessed single-word receptive vocabulary. The Test of Narrative Language (TNL; Gillam & Pearson, 2004) appraised narrative comprehension and production. Part I of the Diagnostic Evaluation of Language Variation-Screening Test (DELV-S; Seymour, Roeper, de Villiers, & de Villiers, 2003) assessed

language variation. Criterion scores from Part I of the DELV-S classified children into one of three language variation categories: *no*, *some*, or *strong* variation from MAE. In addition to diagnostic tests, children participated in a pure tone air conduction hearing screening in which 96.1% passed. The three children who failed the hearing screening performed within normal limits on the norm-referenced tests. Therefore, these children remained in the sample. Table 1 reports descriptive statistics for children's performance on the diagnostic measures.

Narrative Procedures

Elicitation. Children produced spoken narratives elicited from a personal prompt in addition to a fictional prompt from the TNL. To elicit a personal narrative, the second author provided a sample personal story. The examiner played the personal story from a laptop. Then the examiner said to the child, "Now it's your turn to tell me a personal story. You can tell a story about a time when you or someone you know got in trouble, had an accident, had a fun birthday party, was embarrassed or scared, or any topic you choose." When the child looked ready to speak, the examiner then asked, "what topic are you going to tell a story about? Ok, CHILD'S NAME, tell me the best story you can about TOPIC." The examiner used back-channeling (e.g., "mm-hm, yeah") to indicate sustained interest, and asked "Is that all?" or "Was that the end?" when the child appeared to have finished. Fictional narratives were elicited in the context of the TNL, cued by a single picture of visiting aliens; children were instructed to tell a complete story, cued by a single picture of visiting aliens.

Transcription. Undergraduate research assistants and graduate clinicians in communication sciences and disorders completed self-paced online training provided by the SALT website. After practicing on a sample dataset, they then orthographically transcribed audio samples of each narrative. A second undergraduate coder transcribed eight of the 76 narratives

(10%) to establish interrater agreement on SALT transcription (C-unit segmentation, bound morpheme marking, free morpheme transcription). We calculated interrater agreement using Krippendorff's alpha (Krippendorff, 2004).

Krippendorff's alpha agreement levels that are greater than .67 are acceptable for tentative conclusions, and agreement levels that are greater than .80 indicate adequate agreement (Hayes & Krippendorff, 2007). Agreement on occurrences of C-units—independent clauses plus their modifiers—reached .9994 in personal narratives and 1.0 in alien narratives. Agreement on morpheme marking (e.g., *liked runnin*g = gerund vs. *was run/ ing* = progressive) reached .9999 in personal narratives and 1.0 in alien narratives. Agreement on the presence of a word (e.g., *in* vs. *and*) reached .9999 in personal narratives and 1.0 in alien narratives. Thus, all transcription agreement levels were adequate.

Study Variables

Language sample analysis. The analysis set for each transcript was evaluated via Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2012) to obtain values for MLU in words across all C-units and for NDW. NDW was divided by total number of C-units to create NDW rate, which controlled for volubility. NDW rate reports the average number of different words per C-unit, thus distinguishing children who use consistently varied vocabulary from those who produce a large number of utterances.

Rare vocabulary analysis. We assessed the narratives for low-frequency vocabulary using the Wordlist for Expressive Rare Vocabulary Evaluation (WERVE), a list of 2079 words that; can be used within the CLAN software program (MacWhinney, 2000) to generate a tally of the rare vocabulary in a language sample (Mahurin-Smith, DeThorne, & Petrill, 2015). WERVE evaluates the vocabulary in a language sample based on a comparison with a corpus of 1.2

million words used in conversations and narratives by school-aged children Words that appeared 0-15 times in the large corpus are counted as rare. Though WERVE is drawn from language samples produced by a large sample of European American children from middle-income backgrounds (Mahurin-Smith et al., 2015), our prior work shows that African American children perform similarly (Mills et al., 2017).

We used the FREQ command in CLAN to derive type and token tallies for rare vocabulary. We assessed each instance of rare vocabulary use in context. We eliminated transcription artifacts (e.g., a transcriber typed "jargon" as part of an utterance rather than a comment), kinship terms (e.g., Oma or Opa occurred infrequently as labels for grandparents, but they were not treated as rare vocabulary because of the universality of the early-acquired concept they represent), and the like. To be counted, each instance of rare vocabulary needed to be used with an approximation of semantic appropriateness. For instance, the TNL includes a picture of an alien family. A child who describes the alien parents as "two loving couples" is not using "couple" with adult-like semantic accuracy, but at the same time an emerging awareness is clearly evident: the word "couple" can refer to partners who share a romantic attachment. The first author completed initial judgments of semantic appropriateness for all of the items in the CLAN output; a trained research assistant independently assessed all instances of rare vocabulary. Point-to-point reliability was 95.5%; disagreements were reviewed and discussed to establish consensus. To limit confounding with sample length (Greenhalgh & Strong, 2001), we obtained rare-word density (RWD) values by dividing the number of utterances in a story into the number of rare-word types in each story. For example, a child who used 4 instances of rare vocabulary types in a story that was 10 utterances in length would have a rare word density of 0.4, while one instance of rare vocabulary in a 10-utterance story would be recorded as 0.1. The

number of rare-word tokens is reported in Table 2 but was not analyzed statistically; a child's story might include the word "ornithomimus" in 10 separate places, but because of our focus on rare-word types rather than tokens, the word would only be counted once.

Language variation. In addition to their narrative performance, we examined children's language variation to determine the extent to which their dialect differed from MAE using Part I of the DELV-S. Comprising contrastive items related to phonology and morphosyntax, Part I of the DELV-S yields criterion scores that categorized children as speaking with *no*, *some*, or *strong* variation from MAE. We also evaluated language variation by coding African American English features in all transcripts and obtaining Dialect Density Measure per word values based on morphosyntactic features (DDMw; Craig & Washington, 2006). DDM calculations were completed by the second author.

Statistical Analysis

One of the challenges of investigating rare vocabulary is that it is, by definition, a less common occurrence than others observed by language sample researchers. It is not unusual for children to produce stories that rely on common vocabulary, resulting in highly skewed distributions with a number of zeroes. This phenomenon was addressed by using zero-inflated regression models to evaluate the relationships between rare vocabulary use and the other variables of interest (Ferrari & Cribari-Neto, 2004; Long, 1997; Ospina & Ferrari, 2010). A zero-inflated beta regression is a type of generalized linear model suitable for non-normally distributed proportion data with a large number of zeroes. To evaluate whether rare vocabulary was related to language variation, as measured by the DELV-S and DDM, we conducted a Spearman rank order correlational analysis. For Spearman rank order correlations, we report *p* values as well as effect sizes—characterizing them as small (r_s = .10–.29), medium (r_s = .30–.49),

or large (r_s = .50–1.0)—as discussed in Cohen and Cohen (1983). An alpha level of .05 was set for both statistical procedures.

Results

As in previous research into rare vocabulary use, we found that distribution of rare words was highly skewed. This distribution is illustrated in Figure 1. Of the 76 children in the sample, 29 (38.2%) used no rare vocabulary words across both of their stories, and 18 (23.7%) used only one rare word. Of the remaining children, 29 used multiple rare words, with one of these children using 16 different rare words. Across the sample, the mean number of types used was 1.8 (SD = 2.57), while the median was 1. Density values ranged from 0 to 0.16, with a mean of 0.036 (SD = 0.039) and a median of 0.029. Table 2 describes children's performance on narrative measures. NDW rate and MLU were combined into a composite because these two variables were strongly correlated (r = 0.61). We created a composite variable by deriving z scores for each variable and averaging them.

Our first research question considered the associations among rare vocabulary use, standardized vocabulary test results, and language sample measures; our second research question focused on the possible influence of sex on rare vocabulary use. To address research questions 1 and 2, we built a zero-inflated beta regression model that evaluated the relationships among rare vocabulary use, standardized vocabulary test results, and language sample measures while controlling for sex and age. These results are summarized in Table 3. Language sample measures, represented in these models by a composite of MLU and NDW rate, were positively significant (β = 0.13, p < 0.01) in predicting the use of rare vocabulary. Sex also showed a strong relationship (β = 0.29, p < 0.01) with rare vocabulary use, meaning that boys' rare vocabulary density was significantly higher, on average, than girls'. PPVT-4 performance lost statistical

significance in the full model, though it was statistically significant in more parsimonious models that excluded language sample results ($\beta = 0.01$, $p < 0.05$). The full model yielded a modest Cox and Snell pseudo-R^2 of .16. Figure 2 illustrates the relationships across rare vocabulary use, the composite language sample measure, and sex.

Because the models assessing the relationship between rare vocabulary use and sex yielded the unexpected finding of a significant advantage for boys, we conducted a supplemental analysis of the associations between participants' sex and the other language sample variables: the MLU/NDW rate composite as well as the total number of complete utterances. No significant results were obtained via MANOVA, $F(3) = 0.45$, $p > .05$. We also evaluated the potential impact of outliers. As shown in Figure 1, one participant used considerably more rare vocabulary than the others. While the use of density values markedly reduced the skewness of the distribution, we also repeated the analyses without that participant's data. In this subset of the data, male sex remained a statistically significant predictor of rare vocabulary use.

Our third question centered on the relation between children's language variation and rare vocabulary usage. As described previously, we evaluated language variation using the DELV-S coupled with DDMw results for morphosyntactic features. Part I of the DELV-S characterized the sample's language variation as follows: 30.3% no variation from MAE ($n = 23$); 18.4% some variation from MAE ($n = 14$); and 51.3% strong variation from MAE ($n = 39$). Because of the ordinal nature of these language variation scores, we used Spearman rank order correlation analyses to evaluate the relationship between language variation and rare vocabulary use. We found no effect, $\rho = -.04$, $p > .05$. DDMw results ranged from 0 - .083, with a mean of .025 (SD = .019) and a median of .19. Table 4 breaks down these DDMw results with respect to DELV-S classification, illustrating the increase in mean DDMw across the three DELV-S groups. We

assessed the relationship between rare vocabulary density and DDMw with a Pearson correlation. Again we found no effect, r = -.08, $p > .05$.

Discussion

Review of Findings

Our first research question addressed the relationships among rare vocabulary and other language measures, specifically PPVT-4 scores and a composite language sample measure composed of MLU and NDW rate. In the present study, as in prior work on rare vocabulary usage, there was a robust statistical relationship between the composite language sample measure and rare vocabulary use. Previous findings suggested that there might be a weaker association between rare vocabulary use and PPVT-4 scores (DeThorne et al., 2008), and this proved true: PPVT-4 performance predicted rare vocabulary use effectively in isolation but lost significance in the full models. In contrast to our findings for the first research question, our analysis of the second research question yielded surprising results. We expected girls to outperform boys in rare vocabulary production. However, in this sample, school-age boys were more likely than school-age girls to include rare vocabulary in their narratives. Our third research question investigated the relationship between language variation and rare vocabulary use. As we hypothesized based on previous work, there was no association between the two, regardless of whether AAE use was measured via the DELV-S or via DDMw.

Interpretations

In prior work on language samples, other investigators have reported that standardized tests correlate strongly, and that language sample measures correlate strongly, but the two types of measures may correlate more modestly with each other (DeThorne et al., 2008). The present study corroborates this finding. In this investigation, we report that children’s use of rare

vocabulary – a language sample measure – aligns closely with their performance on other language sample measures. In contrast, the association between rare vocabulary use and PPVT-4 results, which was statistically significant in more parsimonious models, lost statistical significance in the full model. Because some norm-referenced tests can underestimate the ability of culturally and linguistically diverse children (Stockman, 2010; Walton & Spencer, 2009), it is important for clinicians to bear in mind that the associations between language samples and formal tests may be variable and that language sampling offers a more ecologically valid context in which to assess language abilities. These findings of a strong association between rare vocabulary use and more traditional language sample measures suggest that it may be a useful adjunct in evaluating semantic skill. High number of total words (NTW) may be associated with volubility as well as strong vocabulary; similarly, high number of different words (NDW) may indicate unusually frequent topic shifts. Assessing for effective deployment of uncommon vocabulary may thus offer useful additional information about a child's semantic skills.

Our theoretical framework for this investigation placed particular emphasis on the role of individual differences in vocabulary acquisition: we point out that children acquire new vocabulary items most readily when the words are especially salient for them personally (Mani & Ackerman, 2018) when their existing semantic networks are densely populated with related terms (Borovsky et al., 2016), and when the task is experienced as intrinsically rewarding (Ripolles et al., 2014). We suggest that for children from culturally and linguistically diverse backgrounds, an assessment framework that accommodates individual differences – and potentially, the use of a tool such as WERVE, which was designed to capture highly varied expressive vocabulary items in naturalistic contexts – can showcase their unique strengths effectively.

Our investigation of RQ1 yielded the hypothesized results, but our findings for RQ2, regarding sex differences in rare vocabulary use, were unexpected. We found an advantage for boys that persisted even after a potential outlier was removed from the dataset. While much of the literature describes a pattern in which girls outperform boys on vocabulary tasks, our findings align with those of Allison et al. (2011), who also reported that the boys in their sample of low-income preschoolers scored higher than the girls by a wide margin (Cohen's d >1). The sparse literature in this area suggests a need for additional research into this question; we note that these findings of sex-related differences in vocabulary *use* do not speak to sex-related differences in vocabulary *size*. If these findings are replicated, they will raise interesting questions in view of the overrepresentation of African American boys from low-income communities on special education caseloads. Is it possible that implicit bias, manifesting itself in lower expectations for African American students (Wood, Kaplan, & McLoyd, 2007), might cause facility with rare vocabulary to go unnoticed and unsupported? If so, identifying individual strengths is a key step toward addressing underachievement. Regardless of whether the pattern of sex differences observed here is noted in other datasets, the narratives under consideration in this study highlight an important trend in semantic development for both boys and girls. The example sentences presented in the Appendix and broken down by sex illustrate the presence of abstract and literate language among a selection of the school-age children in this sample.

In congruence with our prior findings in school-age children, rare vocabulary appeared to be culturally fair in the current study. That is, school-age children with higher DDMw, or strong variation from MAE on the DELV-S, told stories that were as lexically sophisticated as those with lower DDMw and/or DELV-S results indicating little or no variation from MAE. Given that it aligns with norm-referenced measures of language ability, we propose that rare vocabulary is a

culturally fair measure with diagnostic utility. We can only speculate as to why WERVE might provide culturally fair results, but the need to build an increasingly varied semantic network in order to represent an increasingly varied array of underlying concepts is a task that spans cultures and dialects. For some time now, researchers in the discipline of communication sciences and disorders have made the argument that standardized tests alone are insufficient tools for accurate identification of language impairment (Laing & Kamhi, 2003; Ebert & Pham, 2017; Spaulding, Plante, & Farinella, 2006). For children from culturally and linguistically diverse communities, the need for assessment strategies that look beyond standardized test scores remains especially urgent; addressing educational disparities effectively requires that school-based assessment approaches value children's home language (Carter, 2010; Hallett, 2015; Stockman, 2010).

Clinical Significance

For SLPs who provide screenings, evaluations, and intervention for children from low-income communities, the need for culturally fair assessment tools is clear. Ideally, vocabulary intervention will place some emphasis on words that are particularly salient for a given child, and vocabulary evaluation tools will be able to capture these individualized gains. The example sentences in the Appendix provide clear evidence of the interest that rare vocabulary can add to a child's narratives. Moreover, these sentences show the capacity of school-age African American children from low-income communities to produce rich language in the context of narrative generation tasks, and the co-occurrence of rare vocabulary with morphosyntactic features of AAE. Traditionally, vocabulary assessment in school-age children has relied on highly structured standardized testing tasks that feature a highly constrained range of vocabulary. We see potential utility in a generative task that allows children to deploy the specialized vocabulary that is most salient for them, and that avoids the persistent and troubling issue of cultural bias in standardized

tests. SLPs interested in exploring WERVE use can find detailed instructions in the appendix to Mahurin-Smith, DeThorne, & Petrill, 2015.

In light of Common Core State Standards aimed at vocabulary acquisition and use, our findings point to a need to lay the groundwork for children's college and career readiness (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). The domain-specific words that sixth graders are required to use accurately, for example, are often rare. Increasing lexical sophistication, therefore, should be a target that educators and other school-based professionals such as speech-language pathologists aim to support in African American children. Some theoretical and empirical support is available for educators and speech-language pathologists seeking guidance about best practices; Sobel, Cepeda, and Kapler (2011) recommended spacing exposures to unfamiliar vocabulary words one week apart for school-age children. Further research is needed into strategies to facilitate children's learning of rare vocabulary.

Limitations

In this sample we observed a floor effect, with a relatively large proportion of the sample using no rare words. This finding suggests that WERVE's specificity is likely to be higher than its sensitivity in samples of children with language impairment. Additionally, the high number of zeroes observed in this sample suggests that WERVE is most likely to yield meaningful results in language samples from older school-age children. In the present study we mitigated this threat to validity by employing a statistical procedure tailored for datasets that include many zeroes.

Future Directions

Our work demonstrates that rare vocabulary may be a culturally fair measure of narrative language. The presence of rare vocabulary in a child's language sample may offer some

diagnostic utility: children who use rare words in their narratives seldom score outside the normal range on norm-referenced instruments. In future work we plan to address WERVE's utility in identifying diagnostic risk for language concerns in older school-age children. A more dynamic approach to language sample collection might also serve as a strategy for purposeful elicitation of rare vocabulary (e.g., an examiner might say, "You said the aliens were ___. What are some other ways we could say that?"). Unfortunately, screeners of language ability for AAE speakers are lacking, and it is increasingly important for speech-language pathologists to have culturally fair measurement options available for assessing AAE-speaking children who may be at risk for language impairment. Language sample measures, including WERVE, may thus play an increasingly critical role in culturally fair assessments of children who speak AAE.

Future research could address the utility of WERVE in samples of African American children with higher dialect density. Average usage of morphosyntactic features of AAE in this sample fell within the range described by Thompson, Craig, and Washington in their 2004 paper, but the mean DDMw here was lower than the mean reported by those authors. Some of this difference may be explained by the presence of participants in this sample who were not African-American; some of the difference may be related to the school-based context for sample elicitation. It could be valuable, however, to evaluate the use of rare vocabulary within a sample of children with a higher concentration of morphosyntactic features of AAE.

Future research should also address the sex difference observed in this sample, where boys demonstrated more frequent use of rare vocabulary than did girls. Little existing research has considered the interplay between SES and sex differences in language; further investigation of the phenomenon observed here may equip service providers to target children's needs more effectively.

The use of rare vocabulary in written language may also be a fruitful consideration for future researchers. Both the data collected for this project (i.e., spoken narratives) and the tool used to evaluate them (WERVE, developed specifically for analyzing spoken language) render the current dataset unsuitable for this purpose, but school-age children are beginning to demonstrate mastery of increasingly complex written language and their written narratives may be a rich source of information about less common vocabulary items.

Conclusion

Rare vocabulary appears to be a culturally fair measure that aligns with existing language sample measures across a critical age range for vocabulary growth. In this study of children from low-income communities, rare vocabulary proved to be a particular strength for boys, who used significantly more rare vocabulary than the girls in the sample. To adhere to best practices of assessing language in school-age African American children, it is important for school-based service providers to employ measures of narrative language that are both culturally-fair and ecologically valid. Our findings point to rare vocabulary use, which can be assessed rapidly via automated analysis, as a candidate.

Acknowledgments

We are grateful to the parents who allowed their children to participate in this study. We appreciate the school staff who provided time and space for data collection. We extend gratitude to students who assisted with data collection (Aj'a Johnson) and narrative transcription (Julie Applebaum, Taliah Basar, Canaro Cummings, Kaleigh Keplinger, Courtney Winfrey). We thank Monica Fox and Kelsey Osterhage for assistance with establishing reliability, and Elena Pivek for research assistance. We are grateful to Laura DeThorne and Steven Petrill for their collaboration on WERVE with the first author, and to Peter Rankin for his input on zero-inflated models. Finally, we are thankful for a Social and Behavioral Sciences Small Grant, awarded to the second author. This study was approved by the Behavioral and Social Sciences Institutional Review Boards at The Ohio State University and at Illinois State University.

Disclosure statement: The authors have no financial disclosures to report.

References

Allison, C., Robinson, E., Hennington, H., & Bettagere, R. (2011). Performance of low-income African American boys and girls on the PPVT-4: A comparison of receptive vocabulary. *Contemporary Issues in Communication Sciences and Disorders, 38,* 20-25.

Barbu, S., Nardy, A., Chevrot, J.P., Guellaï, B., Glas, L., Juhel, J., & Lemasson, A. (2015). Sex differences in language across early childhood: Family socioeconomic status does not impact boys and girls equally. *Frontiers in Psychology, 6,* 1-10.

Benson, M.S. (1997). Psychological causation and goal-based episodes: Low-income children's emerging narrative skills. *Early Childhood Research Quarterly, 12*, 439-457.

Bitteti, D. & Hammer, C.S. (2016). The home literacy environment and the English narrative development of Spanish-English bilingual children. *Journal of Speech, Language, and Hearing Science, 59,* 1159-1171.

Blake, R. & Cutler, C. (2003). AAE and variation in teachers' attitudes: A question of school philosophy? *Linguistics and Education, 14,* 163–191.

Borovsky, A., Ellis, E.M., Evans, J.L., & Elman, J.L. (2016). Lexical leverage: Category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science, 19*(6), 918-932.

Carter, P.L. Race and cultural flexibility among students in different multiracial schools. *Teachers College Record, 112*, 1529–1574.

Cohen, J. & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed).* New York: Erlbaum.

Connor, M.H., & Boskin, J. (2001). Overrepresentation of bilingual and poor children in special education classes: A continuing problem. *Journal of Children and Poverty, 7*(1), 23-32.

Craig, H.K., Kolenic, G.E., & Hensel, S.L. (2014). African American English-speaking students: A longitudinal examination of style shifting from kindergarten through second grade. *Journal of Speech, Language, and Hearing Research, 57,* 143-157.

Craig, H.K. & Washington, J.A. (2000). An assessment battery for identifying language impairments in African American children. *Journal of Speech, Language, and Hearing Research, 43,* 366-379.

Craig, H. K., & Washington, J. A. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research*.

Craig, H.K., & Washington, J.A. (2006). *Malik goes to school: Examining the language skills of African American students from preschool—5th grade*. Mahwah, NJ: Lawrence Erlbaum Associates.

Craig, H. K., Zhang, L., Hensel, S. L., & Quinn, E. J. (2009). African American English-speaking students: An examination of the relationship between dialect shifting and reading outcomes. *Journal of Speech, Language, and Hearing Research, 52*, 839–855.

DeThorne, L. S., Petrill, S. A., Hart, S. A., Channell, R. W., Campbell, R. J., Deater-Deckard, K., Thompson, L. A., & Vandenbergh, D. J. (2008). Genetic effects on children's conversational language use. *Journal of Speech, Language, and Hearing Research, 51*, 423–435.

Dudley-Marling, C., & Lucas, K. (2009). Pathologizing the language and culture of poor children. *Language Arts, 86,* 362-370.

Fey, M.E., Catts, H.W., Proctor-Williams, K., Tomblin, J.B., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research, 47,* 1301-1318.

Gámez, R.B. & Lesaux, N.K. (2012). The relation between exposure to sophisticated and complex language and early-adolescent English-only and language minority learners' vocabulary. *Child Development, 83*, 1316-1331.

Gillam, R. B. & Pearson, N. (2004). *Test of Narrative Language*. Austin, TX: Pro-Ed.

Greenhalgh, K.S., & Strong, C. J. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, & Hearing Services in Schools, 32*, 114-125.

Hallett, J. (2015) Contexts for student AAE use in the classroom, *Critical Inquiry in Language Studies, 12*, 1-26, DOI: 10.1080/15427587.2015.997648.

Hamilton, R., McCoach, D.B., Tutwiler, S.M., Siegle, D., Gubbins, E.J., Callahan, C.M. Brodersen, A.V., & Mun, R.U.(2018). Disentangling the roles of institutional and individual poverty in the identification of gifted students. *Gifted Child Quarterly, 62*, 6-24.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77-89.

Hendricks, A.E. & Adlof, S.A. (2017). Language assessment with children who speak nonmainstream dialects: Examining the effects of scoring modifications in norm-referenced assessment. *Language, Speech, and Hearing Services in Schools*, *48*, 168-182.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26,* 55-88.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236-248.

Ivy, L. J., & Masterson, J. (2011). A comparison of oral and written English styles in African American students at different stages of writing development. *Language, Speech, and Hearing Services in Schools, 42*, 31–40.

Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and Individual Differences, 27,* 82-89.

Laing, S.P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools, 34,* 44-55.

Leu, D.J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2014). The New Literacies of Online Research and Comprehension: Rethinking the Reading Achievement Gap. *Reading Research Quarterly, 50,* 37-59.

MacWhinney, B. (2000). Computerized language analysis (Version 7.28.09) [Computer software]. Pittsburgh, PA: Author.

Mahurin-Smith, J., Dethorne, L., & Petrill, S. (2015). From aardvark to ziggurat: A new tool for assessing children's use of rare vocabulary. *Clinical Linguistics & Phonetics, 29,* 441-454.

Mani, N., & Ackermann, L. (2018). Why do children learn the words they do? *Child Development Perspectives, 12*(4), 253-257.

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45,* 998-1014..

McKeown, M. G., & Beck, I. L. (2004). Direct and rich vocabulary instruction. In J.F. Baumann & E.J. Kame'enui (Eds.) *Vocabulary instruction: Research to practice*, (pp. 13-27). Guildford Press.

Miller, J. & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT), Research Version 2012 [Computer Software]. Middleton, WI: SALT Software, LLC.

Mills, M.T. and Fox, M. (2016). Language variation and theory of mind in typical development: An exploratory study of school-age African American narrators. *American Journal of Speech-Language Pathology*, *25*, 426-440.

Mills, M.T., Mahurin-Smith, J., & Steele, S.C. (2017). Does rare vocabulary use distinguish giftedness from typical development?: A study of school-age African American narrators. *American Journal of Speech-Language Pathology, 26*, 511-523.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). Common Core Standards. Washington, DC: National Governors Association for Best Practices, Council of Chief State School Officers.

Piotrkowski, C.S., Botsko, M., & Matthews, E. (2000). Parents' and teachers' beliefs about children's school readiness in a high-need community. *Early Childhood Research Quarterly*, *15*, 537-558.

Price, J.R., Roberts, J.E., Jackson, S.C. (2006). Structural development of the fictional narratives of African American preschoolers. *Language, Speech, and Hearing Services in Schools, 37*, 178–190.

Ramirez, S.M., & Stafford, R. (2013). Equal and universal access? Water at mealtimes, inequalities, and the challenges for schools in poor and rural communities. *Journal of Healthcare for the Poor and Underserved, 24*, 885-891.

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H. J., ... & Noesselt, T. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology, 24*(21), 2606-2611

Seymour, H. N., Roeper, T. W., de Villiers, J., & de Villiers, P. A. (2003). Diagnostic evaluation of language variation—Screening test. San Antonio, TX: Pearson.

Snow, C.E., Porche, M.V., Tabors, P.O., & Harris, S.R. (2007). *Is literacy enough? Pathways to academic success for adolescents.* Baltimore, MD: Brookes.

Sobel, H.S., Cepeda, N.J., & Kapler, I.V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology, 25,* 763-767.

Spaulding, T.J., Plante, E., & Farinella, K.A. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech, and Hearing Services in Schools, 37,* 61-72.

Stockman, I.J. (2010). A review of developmental and applied language research on African American children: From a deficit to difference perspective on dialect differences. *Language, Speech, and Hearing Services in Schools, 41,* 23-38.

Sullivan, A.L., & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. *Exceptional Children, 79,* 475-494.

Uccelli, P. & Páez, M.M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish Skills. *Language, Speech, and Hearing Services in Schools, 38*, 225-236.

Walton, G.M., & Spencer, S.J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science, 20*(9), 1132-1139.

Wheeler, R. (2019). Attitude change is not enough: Disrupting deficit grading practices to disrupt dialect prejudice. *Proceedings of the Linguistic Society of America, 4,* 1-12.

Wheeler, R., Cartwright, K.B., & Swords, R. (2012). Factoring AAVE into reading assessment and instruction. *The Reading Teacher, 65,* 416-425.

Wolfram, W. & Schilling-Estes, N. (2006). *American English: Dialects and Variation,* 2nd ed. Malden, MA: Blackwell.

Wood, D., Kaplan, R., & McLoyd, V.C. (2007). Gender differences in the educational expectations of urban, low-income African American youth: The role of parents and the school. *Journal of Youth and Adolescence, 36,* 417-427.

Figure 1. Distribution of rare vocabulary types.

Figure 2. Regression model of the relationships between language sample measures and rare vocabulary use.

Appendix

Sample Sentences Including Rare Vocabulary

Personal Narrative Samples

They kept **investigating**

We had a **discussion** about where she was going.

They told my mom that I had to have **surgery.**

We was in **separate** classes.

We had to wait a couple of minutes before we could get back in the pool so our food could **digest**.

He **smeared** it on his face.

My parents **panicked.**

We went back to **shore.**

He **lowered** the **price** for us.

Alien Narrative Samples

I thought they were a **myth.**

But Brandon was scared, like a **coward.**

He don't know what they're **capable** of.

One day, Jon and Mary were going to an alien **convention.**

After **inspecting** some of their alien **gadgets** and what-not, they went back and took the alien **device**,that the aliens gave them, back to their home.

There was a strange **satellite** that came out of the sky.

Lily wanted to go out and **introduce** them.

Get off my **lawn**.

Figure 1

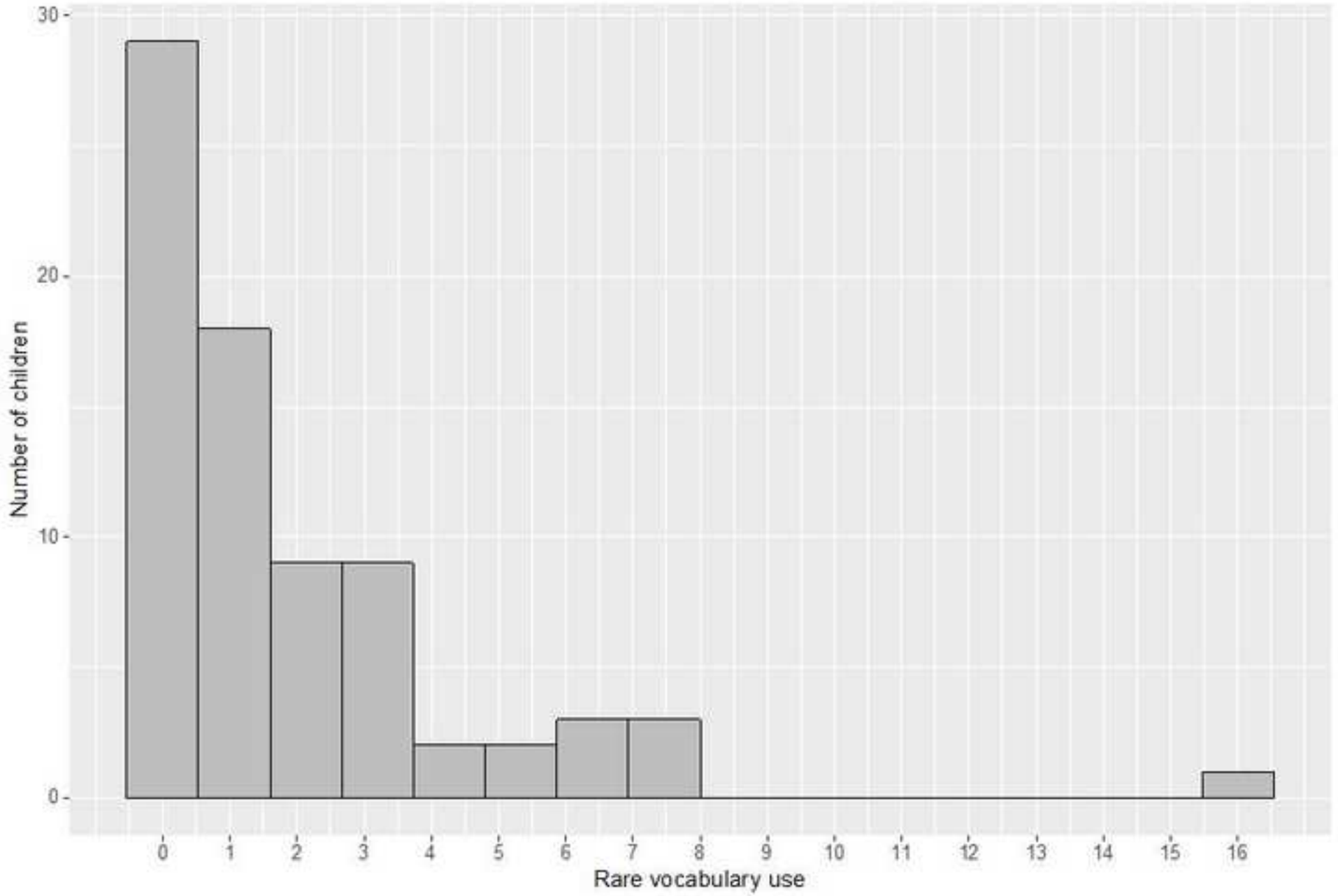

Figure 2

Click here to access/download;Figure;Figure2 revised.png

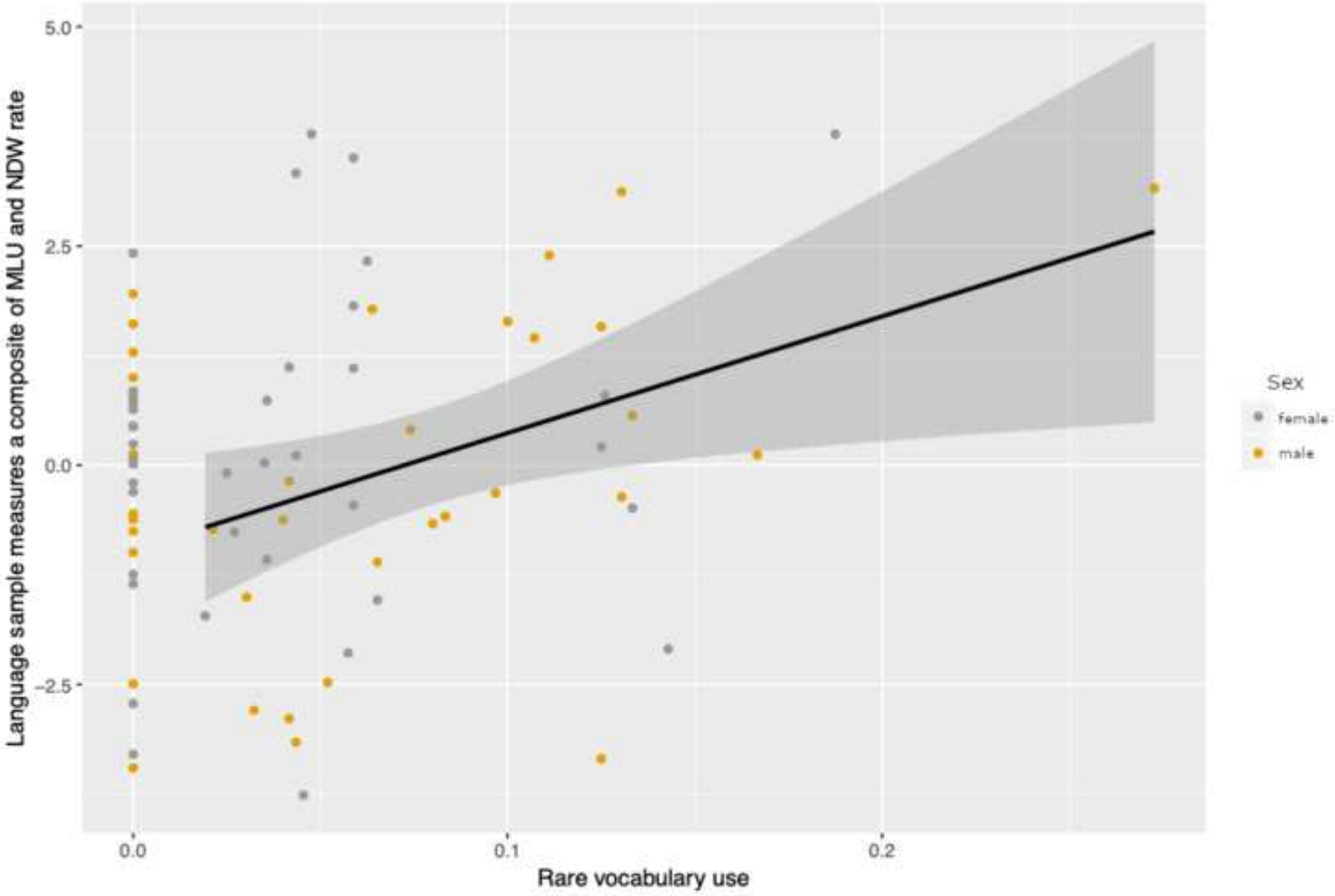

Table 1. Mean performance on standardized tests (n = 76).

Measure	Mean	Standard Deviation	Min	Max	Range
PPVT-4	94.08	11.83	72	123	51
TNL	96.17	12.51	64	136	72

Note. PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007); TNL = Test of Narrative Language (Gillam & Pearson, 2004). The PPVT-4 and TNL have a mean of 100 and standard deviation of 15.

Table 2. Mean performance on narratives (*n* = 76).

Variable	Mean	Median	SD	Min	Max
MLU	8.55	8.43	1.42	4.66	11.72
NDW rate	4.25	4.20	0.88	2.21	6.56
TNCU	32.3	24	25.3	9	154
TNW	258.3	207	218.2	52	1166
RWD	.049	.038	0.06	0	0.27
RWT	1.80	1.00	2.57	0	16

Note. Study variables from personal- and alien stories were pooled. MLU = mean length of communication units; NDW rate = number of different word roots divided by the total number of communication units; TNCU = total number of C-units; TNW = total number of words; RWD = rare-word density; RWT = rare word tokens.

Table 3. Results from zero-inflated models.

Predictor variable	Coefficient	Std. Error	*p* value
Chronological age	-0.11	0.08	.12
Sex (male)	0.44	0.15	.007**
PPVT-4	0.01	0.01	.14
MLU/NDW rate	0.13	0.04	.003**

Note: - indicates $p < .10$, * indicates $p < .05$, ** indicates $p < .01$.

Table 3. Dialect Density Measure in Words (DDMw) values as a function of results on the Diagnostic Evaluation of Language Variation (DELV).

DELV Classification	DDMw Range	DDMw Mean (SD)
No variation (*n* = 23)	0-.06	.016 (.015)
Some variation (*n* = 14)	0-.05	.020 (.012)
Strong variation (*n* = 39)	.005-.083	.032 (.020)

Appendix

Sample Sentences Including Rare Vocabulary

Personal Narrative Samples

They kept **investigating**

We had a **discussion** about where she was going.

They told my mom that I had to have **surgery.**

We was in **separate** classes.

We had to wait a couple of minutes before we could get back in the pool so our food could **digest**.

He **smeared** it on his face.

My parents **panicked.**

We went back to **shore.**

He **lowered** the **price** for us.

Alien Narrative Samples

I thought they were a **myth.**

But Brandon was scared, like a **coward.**

He don't know what they're **capable** of.

One day, Jon and Mary were going to an alien **convention.**

After **inspecting** some of their alien **gadgets** and what-not, they went back and took the alien **device**,that the aliens gave them, back to their home.

There was a strange **satellite** that came out of the sky.

Lily wanted to go out and **introduce** them.

Get off my **lawn**.