LSHSS

Research Article

The Effects of Visual Stimuli on the Spoken Narrative Performance of School-Age African American Children

Monique T. Mills[a]

Purpose: This study investigated the fictional narrative performance of school-age African American children across 3 elicitation contexts that differed in the type of visual stimulus presented.
Method: A total of 54 children in Grades 2 through 5 produced narratives across 3 different visual conditions: no visual, picture sequence, and single picture. Narratives were examined for visual condition differences in expressive elaboration rate, number of different word roots (NDW) rate, mean length of utterance in words, and dialect density. The relationship between diagnostic risk for language impairment and narrative variables was explored.
Results: Expressive elaboration rate and mean length of utterance in words were higher in the no-visual condition than in either the picture-sequence or the single-picture conditions. NDW rate was higher in the no-visual and picture-sequence conditions than in the single-picture condition. Dialect density performance across visual context depended on the child's grade, so that younger children produced a higher rate of African American English in the no-visual condition than did older children. Diagnostic risk was related to NDW rate and dialect density measure.
Conclusion: The results suggest the need for narrative elicitation contexts that include verbal as well as visual tasks to fully describe the narrative performance of school-age African American children with typical development.

Assessments of spoken narratives have been recommended by many researchers as a more naturalistic and culture-fair assessment than standardized tests for culturally and linguistically diverse children (Bliss, Covington, & McCabe, 1999; Justice et al., 2006; Katz & Champion, 2008). However, narrative assessments still may provide an incomplete picture of children's ability if language variation, through measures of dialect density, is not taken into account as part of narrative analyses (Craig & Washington, 2006; Ivy & Masterson, 2011; Mills, Watkins, & Washington, 2013; Ross, Oetting, & Stapleton, 2004; Schachter & Craig, 2013; N. P. Terry, Mills, Bingham, Mansour, & Marencin, 2013).

The structure of narrative language traditionally has been characterized at two levels: macrostructure and microstructure. *Macrostructure* references the beginning of the story, where the setting and characters are introduced; the middle of the story, where characters encounter conflicts or tensions, and the end of the story, where these problems are resolved (Hughes, McGillivray, & Schmidek, 1997; Justice et al., 2006). The quality of narrative macrostructure has been captured by measures such as story grammar (Bates, 1991; Epstein & Phillips, 2009; Mandler & Johnson, 1977; Merritt & Liles, 1989; Schneider, 1996; Schneider & Dubé, 1997, 2005; Stein & Glenn, 1979), episodic structure (M. S. Allen, Kertoy, Sherblom, & Pettit, 1994; Celinska, 2009), high point analysis (Champion, 2003; Labov, 1972; Peterson & McCabe, 1983), and expressive elaboration (Glenn-Applegate, Breit-Smith, Justice, & Piasta, 2010; Mills et al., 2013; Ukrainetz & Gillam, 2009; Ukrainetz et al., 2005).

In contrast, narrative microstructure references the way that words and sentences work together to build a cohesive story (Hughes et al., 1997). The quality of narrative microstructure traditionally is determined on the basis of semantic and syntactic productivity (e.g., total number of words, total number of utterances), complexity (e.g., number of different words, mean length of utterance), and accuracy (e.g., percentage grammatical utterances). Although macrostructure and microstructure both contribute to a narrative that is well crafted, some researchers argue that microstructure may be a better index of narrative performance

[a]The Ohio State University, Columbus
Correspondence to Monique T. Mills: mills.298@osu.edu
Editor: Marilyn Nippold
Associate Editor: Sonja Pruitt-Lord
Received July 13, 2014
Revision received December 13, 2014
Accepted May 29, 2015
DOI: 10.1044/2015_LSHSS-14-0070

Disclosure: The author has declared that no competing interests existed at the time of publication.

than macrostructure because microstructure is more sensitive to developmental differences between children with language impairment (LI) and those with typical development (TD; Liles, Duffy, Merritt, & Purcell, 1995) and because microstructure may be more dialect neutral (de Villiers, 2004) than macrostructure.

Macrostructure, microstructure, and dialect density represent central aspects of narrative language to assess in school-age African American children. However, it is also important to know how performance on these measures may be influenced by factors that are internal and external to the child. Studies indicate that narrative language improves with age (Burns, de Villiers, Pearson, & Champion, 2012; Ukrainetz & Gillam, 2009) and that narrative language is better formed without a visual stimulus than with a visual stimulus (Gazella & Stockman, 2003; Schneider, 1996; Schneider & Dubé, 2005). The present study builds on previous findings with preliminary examinations of the grade level and visual stimuli effects on narrative macrostructure and microstructure; it extends previous research by also exploring grade and visual stimuli effects on dialect density in a sample of African American children. Knowing whether visual cues affect narrative performance will help improve the ability to accurately assess the narrative language of African American children—a group that may present with narrative language that varies from classroom expectations (Curenton, 2006; Heath, 1983; Michaels, 1981; Weddington, 2010).

Visual Stimuli and Narrative Language

Studies of school-age children with and without LI have elicited spoken narratives using visual stimuli such as films (Hicks, 1991; Scott & Windsor, 2000), pictures in wordless books (Gorman, Fiestas, Peña, & Clark, 2011), picture sequences (Fey et al., 2004), and single pictures (Justice et al., 2006). Clinicians routinely use visual stimuli in the assessment and treatment of narrative language, presumably because they believe that these pictorial materials will aid narrative comprehension and production. Nonetheless, studies have not consistently supported a visual over auditory presentation effect on narrative performance in either preschool-age children (Gazella & Stockman, 2003; Pratt & MacKenzie-Keating, 1985) or school-age children (Masterson & Kamhi, 1991; Schneider, 1996).

The effects of visual stimuli on narrative performance have been explained in terms of the relative task demands on attention, memory (Schneider, 1996; Schneider & Dubé, 1997, 2005), and language (Bates, 1991; Nurss & Hough, 1985; Spinillo & Pinto, 1994) between retelling a narrative that is presented verbally and generating a narrative from a visual stimulus. Comprehending a narrative that is presented verbally requires children to attend to story units and then store them in short-term memory long enough to process them, integrate them with incoming story units, and relate them to an existing story schema—a set of expectations about the internal structure of stories (Mandler & Johnson, 1977; Stein & Glenn, 1979)—which is presumably stored in long-term memory. To retell a narrative that is verbally presented requires children to reformulate the main idea of the story from memory while maintaining the same language structuring—words and sentences used for character and plot development and for temporal ordering. To generate a narrative that is visually presented requires children to extract the gist of the story from the pictures and to develop the language structuring. Thus, the main difference between narratives that are presented verbally and those that are presented visually is the presence or absence of prior language structuring. Given that visual stimuli do not provide language structuring, children may produce macrostructure and microstructure that are less well crafted in elicitation contexts with visual stimuli than without visual stimuli.

Macrostructure

In studies of school-age children from diverse cultural and linguistic backgrounds, visual stimuli effects on narrative macrostructure have been assessed primarily by scoring story grammar units such as setting, character introduction, initiating event, development of plot, resolution, and conclusion. These studies have indicated that school-age children appear to produce written (Bates, 1991) and spoken (Nurss & Hough, 1985; Schneider, 1996; Schneider & Dubé, 2005; Spinillo & Pinto, 1994) narratives with more complete macrostructure without visual stimuli than with visual stimuli. For example, Bates (1991) examined written narrative macrostructure in 35 children with TD, ages 6 to 8 years, in English as a Foreign Language and English as a Second Language classrooms in the United Kingdom. Children produced written narratives across three visual conditions: a single picture, a picture sequence, and no visual stimulus. Results indicated that children wrote better narratives without a visual stimulus than with either a picture sequence or a single picture. Children produced, on average, four story grammar elements in the no-visual condition and only two story grammar elements in both the single-picture and picture-sequence conditions.

Studies of spoken narration in school-age children with TD also suggest that macrostructure is more complete when narratives are elicited without visual stimuli than with visual stimuli. For example, Spinillo and Pinto (1994) examined spoken narrative generation in 60 British and 60 Italian children with TD, ages 4, 6, and 8 years, across four narrative conditions: single picture, picture sequence, no visual, and dictation. Results indicated that narrative differences existed between the visual stimuli conditions (single picture, picture sequence) and the no-visual stimuli conditions (no visual, dictation). Children across age groups told more coherent stories without a visual stimulus than with a visual stimulus, leading the authors to conclude that visual stimuli may mask narrative capabilities.

Nurss and Hough (1985) found that narrative macrostructure was more complete when school-age children narrated without a visual stimulus than with a picture sequence. In their study of 48 children with TD (63% African American) in kindergarten to Grade 3, Nurss and Hough

(1985) indicated that children included a greater number of story grammar elements when narrating an original story (no visual) or narrating a wordless book than when narrating from a picture sequence.

Narrative macrostructure also appears to be better when elicited without visual stimuli than with visual stimuli in school-age children with LI. For example, Schneider (1996) examined the spoken narrative retellings of 16 White Canadian children with LI, ages 5 to 9 years, across four visual conditions: pictures only, oral followed by pictures, oral with pictures, and oral only. Results indicated that children produced the greatest number of story grammar units in the oral-only condition and the smallest number of story grammar units in the pictures-only condition.

Taken together, studies of school-age children indicate a narrative macrostructure advantage in elicitation contexts that do not include a visual stimulus over those that do include a visual stimulus. However, visual stimuli effects on narrative macrostructure must be considered in light of children's grade level. For example, in their study of the spoken narrative retellings of 44 White Canadian children with TD in kindergarten and Grade 2, ages 5 to 7 years, Schneider and Dubé (2005) indicated an interaction between grade and visual condition so that second graders produced a greater number of story grammar units than kindergartners in the oral-with-pictures and oral-only conditions than in the pictures-only condition. Therefore, the current study investigated the grade by visual stimuli interaction effect on narrative macrostructure as indexed by expressive elaboration (EE).

EE has also been the macrostructure variable of interest in other studies that included school-age African American children (Glenn-Applegate et al., 2010; Mills et al., 2013) and in studies of school-age children narrating in response to the visual stimuli provided by the Test of Narrative Language (TNL; Gillam & Pearson, 2004; Ukrainetz & Gillam, 2009). Like story grammar, EE takes into account the overall hierarchal organization of their narratives. EE analysis measures three categories: orientations, appendages, and evaluations. These categories are designed to appraise a child's ability to introduce the setting and characters in the story (orientations); to provide linguistic markers indicating how the story starts, how the plot develops, and how it is resolved (appendages); and to make the story events more intense by using interesting modifiers to describe the internal states and actions of the characters (evaluations).

Microstructure

Studies of microstructure productivity indicate that school-age children with TD present with poorer narrative performance in elicitation contexts with visual stimuli than in those without visual stimuli. For example, Nurss and Hough (1985) found that children produced a greater number of words and a higher descriptive word ratio when telling an original story with no visual stimulus than when narrating in response to either a wordless book or a single picture. In contrast to Nurss and Hough (1985), Schneider (1996) indicated that microstructure productivity—as measured by the average length of utterance, number of different words, and number of utterances—did not differ across visual context. This difference across studies may stem from sample differences between groups of school-age children with LI (Schneider, 1996) and school-age children with TD (Nurss & Hough, 1985).

Extending previous findings to a group of African American school-age children with TD, the current study examined microstructure productivity across visual context using number of different word roots (NDW) rate and mean length of utterance in words (MLU-w). In studies of school-age children, NDW has been shown to increase with age (Fey et al., 2004), to distinguish children with TD from children with LI (Scott & Windsor, 2000), and to improve with intervention (M. M. Allen, Ukrainetz, & Carswell, 2012). MLU-w has been deemed an appropriate assessment of syntactic complexity in school-age children (Hoffman & Gillam, 2004), including those who are African American (Craig, Washington, & Thompson-Porter, 1998).

Dialect Density

Although research is emerging to suggest that narrative analyses of macrostructure (e.g., evaluative language) and microstructure (e.g., cohesion) are dialect neutral and thus appropriate for the assessment of children who speak African American English (AAE; Burns et al., 2012), few studies of the impact of visual stimulus type on the structure of narrative language have been conducted. Only one study of school-age AAE-speaking children was found in which visual stimulus type was varied across language elicitation context; however, this study did not directly appraise the influence of visual stimuli on narrative performance (Thompson, Craig, & Washington, 2004).

Thompson et al. (2004) examined *dialect density*—the extent to which speakers produce linguistic features of AAE—across spoken and literacy contexts in 50 African American third graders with TD. Children described a single picture in the spoken context. In the literacy contexts, children wrote a story on a topic of their choosing and read aloud one-paragraph stories from the Gray Oral Reading Test–Third Edition (Wiederholt & Bryant, 1992). Results indicated that AAE production—as indexed by a dialect density measure (DDM)—was greater when children narrated from a single picture stimulus (spoken narration) than when they narrated from no visual stimulus (narrative writing) or narrated from a written stimulus (oral narrative reading). The pattern of dialect density displayed by children in the Thompson et al. (2004) study may be related to the higher demand for Mainstream American English (MAE) in literacy contexts that restricted AAE production in narrative writing and reading tasks relative to AAE production in the narrative telling task.

Because in the United States AAE is generally assigned lower prestige than MAE, the dialect of wider communication (Adger, Wolfram, & Christian, 2007), its use in "official" settings such as classrooms may not meet teacher expectations. In addition, AAE shares some grammatical features (e.g., zero past tense *–ed*) with LI (Pruitt & Oetting, 2009).

Given that AAE is considered a low-prestige dialect and that it shares linguistic features characteristic of LI, its production may be misinterpreted by school staff as erroneous or even pathological (Curenton, 2006). Therefore, it is important to know whether diagnostic risk for LI is associated with dialect density in school-age children. By examining the relationship between diagnostic risk and narrative language in a sample of African American children with TD, this study provides school-based professionals with expectations for potential linguistic vulnerabilities or strengths that may go undetected by standardized assessment batteries.

Current Study

The present study investigated the effect of visual stimuli on the narrative language—macrostructure, microstructure, and dialect density—produced by school-age African American children. EE was used to measure macrostructure. Two variables—NDW rate and MLU-w—were used to assess microstructure. Macrostructure and microstructure were appraised in fictional narratives because prior studies of the effect of visual stimuli on narrative performance elicited fictional narratives (Bates, 1991; Nurss & Hough, 1985; Schneider, 1996; Schneider & Dubé, 1997, 2005; Spinillo & Pinto, 1994). Further, fictional narratives have been shown to be a particularly informative context for examining structural as well as dialectal aspects of narration in school-age African American children (Mills, 2015; Mills et al., 2013).

The present study addressed (a) whether there are differences related to visual stimulus condition (no visual, picture sequence, single picture) and narrative language (EE rate, NDW rate, MLU-w, and dialect density), (b) whether visual stimulus condition affects the narrative language of children in lower grades (Grades 2 and 3) differently than children in higher grades (Grades 4 and 5), and (c) the relationship between diagnostic risk for LI and narrative language.

On the basis of previous work on school-age children with TD, it was hypothesized that narrative macrostructure performance would be better in response to the no-visual condition than to either the picture-sequence or single-picture conditions (Bates, 1991; Spinillo & Pinto, 1994). On the basis of Nurss and Hough (1985), narrative microstructure was expected to be higher in the no-visual condition than in the picture-sequence condition. Given that no prior study has examined visual effects on dialect density, there was not a firm empirical basis for testing a hypothesis. However, it was conjectured that children may interpret the visual stimuli as a literate context rather than an oral context (Thompson et al., 2004). Hence, children would produce lower dialect density in visual stimuli conditions because these literate contexts would have a higher demand for MAE than the no-visual condition.

Given findings from Schneider and Dubé (1997, 2005), a Grade × Visual Condition interaction was expected so that children in higher grades would produce a more complete macrostructure and microstructure than children in lower grades in the no-visual context compared with the picture-sequence condition. A Grade × Visual Condition interaction on dialect density was not expected to be found because previous studies indicated that the morphosyntactic features of AAE tend not to increase during Grades 2 through 5 (Craig & Washington, 2004: Craig, Zhang, Hensel, & Quinn, 2009).

Last, a negative correlation was expected between diagnostic risk for LI and narrative performance in macrostructure and microstructure on the basis of empirical work, suggesting that narrative language distinguishes school-age children with LI from children with TD (Schneider, 1996; Scott & Windsor, 2000). In contrast, a positive correlation was expected between diagnostic risk for LI and dialect density because African American children with better language flexibility and awareness of the need to shift to school expectations for MAE will be at lower risk for LI than African American children who have not lowered their dialect density during the school-age years. The present hypothesis stems from empirical work suggesting that greater proficiency with MAE tends to result in better performance on academic outcomes in literacy (Craig & Washington, 2004; Craig et al., 2009; Ivy & Masterson, 2011).

Method

Participants

The present study was part of a larger study aimed at identifying school-age African American children with exceptional language ability and differentiating them from African American children with TD and LI. A convenient sample of children was recruited from a public elementary school in central Ohio. Informed consent was obtained from parents or guardians of each child. Children were recruited through parent packets that were sent home with children. They received one of the following two incentive packages: (a) a children's book and $20 or (b) a children's book, $10, and a report of performance on diagnostic assessments. All children received small prizes (e.g., decorative pencils, erasers, and stickers) in each session. Families returned consent forms along with a demographic survey in which one parent provided descriptive information about family socioeconomic status (SES) and the child's educational placement history to characterize the sample.

Of the 75 children who participated in the study, 21 were excluded from the final participant pool because of the following: They were not in Grades 2 through 5 ($n = 9$), they were not African American ($n = 4$), they were English-language learners ($n = 3$), they scored 1.5 *SD* below the mean on the cognitive assessment ($n = 1$), they scored 1.5 *SD* below the mean on the two language assessments ($n = 2$), or they dropped out of the study because of a family move ($n = 2$). Thus, a total of 54 children (29 girls, 25 boys) in Grades 2 through 5 (ages 86–133 months) provided data for this study. There were 28 children in second grade (ages 86–109 months), seven in third grade (ages 94–113 months), 16 in fourth grade (ages 105–133 months), and three in fifth grade (ages 115–133 months). Given the small number of third and fifth graders, two grade groups were created: lower (Grades 2 and 3) and upper (Grades 4 and 5). This grade grouping is congruent with the way that elementary classrooms are often organized developmentally.

SES data were gathered to further describe the sample. Information about family SES and the child's special education placement was derived from the parent survey, which was completed by one of each child's parents. A total of 61% (n = 33) of parent surveys were returned. To index family SES, parents indicated their highest level of education: 36% (n = 12) graduated college, 27% (n = 9) had more than 12 years of education but did not graduate college, 30% (n = 10) graduated high school, and 6% (n = 2) received fewer than 12 years of education. School SES—the percentage of children receiving free or reduced lunch—was used to index family SES for the 39% (n = 21) of parent surveys that were not returned. These 21 parents had children who were being educated in schools in which 100% of children received free or reduced lunch. Thus, the study sample included a greater percentage of children from the lower end of the SES spectrum (61%) than from the higher end of the SES spectrum (39%).

In terms of special education placement, 7% of parents (n = 4) reported that their child had a history of special education services, including speech and language (n = 1), reading (n = 2), and gifted or talented (n = 1). One child had been dismissed from reading services at the time of the study, but the other three were still receiving special education services. These children were included because their performance on tests of cognition, language, and narration were within typical limits.

Procedure

Diagnostic Assessments

Two graduate clinicians and the author administered tests to characterize the overall cognitive, language, and narrative abilities of each child. All testing occurred in two semi-private rooms in a local elementary school. The Test of Nonverbal Intelligence–Fourth Edition (Brown, Sherbenou, & Johnsen, 2012) measured nonverbal cognitive ability. The Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007) assessed single-word receptive vocabulary. The TNL appraised narrative comprehension and production.

To characterize dialect use and diagnostic risk of LI in this sample of children, the Diagnostic Evaluation of Language Variation–Screening Test (DELV-S; Seymour, Roeper, de Villiers, & de Villiers, 2003) was administered. Part I of the DELV-S has criterion scores that categorize children, ages 4 to 12 years, as speaking with no, some, or strong variation from MAE. In this sample, 35% (n = 19) of the children produced no variation from MAE, 11% (n = 6) produced some variation from MAE, and 54% (n = 29) produced strong variation from MAE.

Part II of the DELV-S has diagnostic scores that classify children, ages 4 to 9 years, as being in the lowest risk, low-to-medium risk, medium-to-high risk, or highest risk category for LI. Given the small number of participants in the highest risk category (n = 6), the highest risk category was combined with the medium-to-high risk category to create a total of three diagnostic risk groups: low, medium, and high. Each group included one or more DELV-S risk categories as follows: low (lowest risk), medium (low to medium risk), and high (medium to high risk, highest risk). Forty-five percent of children (n = 21) were in the low-risk group, 23% (n = 11) were in the medium-risk group, and 32% (n = 15) were in the high-risk group.

A one-way analysis of variance (ANOVA) compared the performance of children across the three diagnostic risk groups on standardized tests. Results indicated that children did not differ on Test of Nonverbal Intelligence–Fourth Edition scores, $F(2, 46) = 1.16$, $p = .32$, $\eta^2 = .05$; or Peabody Picture Vocabulary Test–Fourth Edition scores, $F(2, 46) = 2.74$, $p = .07$, $\eta^2 = .11$. There was a main effect for TNL scores, $F(2, 46) = 3.88$, $p = .02$, $\eta^2 = .14$. Post hoc comparison using the Bonferroni test indicated that the low-risk group (M = 103, SD = 13.92) scored significantly higher than the high-risk group (M = 92, SD = 10.24) on the TNL. Table 1 describes the mean performance on the cognitive and language assessments across the three diagnostic risk groups.

Narrative Elicitation

Three fictional narratives were obtained from each child using the TNL. Elicitation followed standard TNL administration procedures except that all narratives were prerecorded and played from a laptop so that each child heard the same stories. The TNL prompt for this story requires the examiner to say to the child, "Try to say the story the same way I said it." Therefore, the directions for the first narrative task (McDonald's) were modified so that the examiner was required to say, "Try to say the story the same way you heard it."

In the first narrative task on the TNL, children hear with no accompanying visual stimulus a story about a family that travels to McDonald's for dinner but fails to bring money to pay for their food (the "McDonald's" story). The child then answers a series of questions about the McDonald's story before retelling the story. In the second narrative task on the TNL, children look at a five-scene picture sequence while listening to the depicted story about a girl who ruins her science project on the way to school ("Shipwreck"). After the child answers a series of comprehension questions about the Shipwreck story, the child tells the story of a little boy who is late for school ("Late") while viewing a different five-scene picture sequence. In the third TNL task, children view a detailed picture while listening to the illustrated story of two siblings who find a treasure chest that is guarded by a dragon ("Dragon"). The child answers a series of questions about the Dragon story. Then, while viewing a different detailed picture, the child tells the story of two siblings who encounter a family of aliens in a park ("Alien"). Each child's narratives were recorded using a digital audio recorder.

Transcription

Spoken narratives were orthographically transcribed using Systematic Analysis of Language Transcripts (Miller & Iglesias, 2010). Graduate research assistants in communication

Table 1. Mean (*SD*) performance on cognitive and language tests of children across grade and diagnostic risk groups.

Grade	Diagnostic risk	TONI-4	PPVT-4	TNL
Lower (Grades 2 and 3)	Low	98 (9.67)	102 (9.53)	105 (15.03)
	Medium	103 (7.13)	95 (9.75)	94 (11.61)
	High	97 (9.41)	92 (13.24)	92 (9.20)
Upper (Grades 4 and 5)	Low	98 (8.64)	94 (8.24)	97 (10.10)
	Medium	100 (10.26)	86 (10.26)	91 (12.00)
	High	97 (4.24)	89 (10.60)	91 (21.21)
All		99 (8.72)	95 (11.24)	97 (12.99)

Note. Standard scores and standard deviations for each test are reported. Diagnostic risk is based on the Diagnostic Evaluation of Language Variation–Screening Test. TONI-4 = Test of Nonverbal Intelligence–Fourth Edition; PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition; TNL = Test of Narrative Language.

sciences and disorders segmented the narratives into communication units (C-units) using Loban's (1976) scoring criteria. *C-units* are independent clauses plus their modifiers, including one main clause along with accompanying subordinate clauses. C-unit segmentation has been established as an appropriate procedure for examining oral language samples (Loban, 1976) and has been used in previous studies on the narrative abilities of African American children (Craig et al., 1998; Hester, 2010; Ivy & Masterson, 2011; Thompson et al., 2004). Labov's (1972) definition of a *minimal narrative* is one that is at least two C-units in length. The children in the current study produced McDonald's narratives that were from five to 19 C-units in length (M = 11, SD = 3.30), Late narratives that were from five to 21 C-units in length (M = 11, SD = 3.74), and Alien narratives that were from five to 62 C-units in length (M = 16, SD = 10.35). Spoken narratives were coded for EE, lexical diversity, syntactic complexity, and dialect density, as outlined in the following sections.

EE rate. The EE analysis followed procedures outlined by Ukrainetz and Gillam (2009). EE analysis consisted of coding words or utterances for 14 elements within the three aforementioned categories: *Appendages* consist of introducer, abstract, theme, coda, and ender; *orientations* consist of name, relation, external, and personality; and *evaluations* consist of modifier, expression, repetition, internal state words, and dialogue. Codable words or utterances were examined in each C-unit. For instance, the utterance "The story is about a sleepy boy who was late for school" would be coded as abstract (appendages) because it summarizes the story before the story begins. The word *sleepy* would be coded as modifier (evaluations) because it is an adjective describing the word *boy*. Words and utterances could receive only one within-category code (e.g., appendages) but multiple between-category codes (e.g., appendages, evaluations). For efficiency, each occurrence of a specific element was coded only on the first mention.

Appendages (except theme) and orientations were awarded 0 or 1 point. Evaluations and theme received 0 or 1 point for up to two occurrences because these elements were expected to occur multiple times throughout the story. The 20 possible points for EE were derived from six possible points from appendages, four from orientation, and 10 from evaluation. Appendix A displays the EE scoring system. Given that the three narratives differed in length, an EE rate was calculated by dividing total EE by the total number of utterances for the McDonald's, Late, and Alien narratives.

Lexical diversity. The NDW—a measure of lexical diversity—is the total number of different words without inflection in a narrative. Children in the current study produced McDonald's stories that contained 27 to 81 different word roots (M = 54, SD = 14.13), Late stories that contained 25 to 81 different word roots (M = 45, SD = 12.68), and Alien stories that contained 22 to 163 different word roots (M = 57, SD = 26.81). Following the procedures outlined in Greenhalgh and Strong (2001), an NDW rate measure was created to control for differences in narrative length among the three visual stimuli conditions by dividing NDW by total number of C-units.

Syntactic complexity. The Systematic Analysis of Language Transcripts program calculated MLU-w to measure children's average sentence length, which is considered a proxy for sentence complexity (Craig et al., 1998). In the current study, children's MLW-w production in the McDonald's stories ranged from 5.13 to 13.10 (M = 8.58, SD = 1.54); in the Late for School narratives, from 4.33 to 11.50 (M = 7.49, SD = 1.60); and in the Alien stories, from 4.57 to 12.14 (M = 7.47, SD = 1.73).

Dialect density. To determine the extent to which child narrators produced AAE, the DDM (Craig & Washington, 2006) was calculated by dividing the total number of morphosyntactic tokens by the total number of words in each narrative, following the procedures of Craig and Washington (2004). For example, four tokens of AAE feature production divided by 100 words would yield a DDM of .04, corresponding to the production of one feature every 25 words. Mean DDM rate was .03 (SD = 0.02) in the current study, corresponding to the production of one AAE feature every 33 words. DDM rates for the children in the current study ranged from .00 to .08, which indicate low to high production of the morphosyntactic features of AAE (Craig & Washington, 2004). In addition, DDM scores were positively correlated with scores on Part I of the DELV-S (ρ = .48, p < .01) in the current sample. Recall that Part I of the DELV-S indicated that 65% of children in the current study produced some (11%) or strong (54%) variation from MAE.

Interrater Agreement

Transcription. Blind interrater agreement was conducted by a second scorer on 36 of the 171 narratives (21%) from a random selection of four children from all three time periods. Point-to-point comparisons determined concordance on C-unit segmentation, bound morpheme marking, and word transcription. These comparisons were calculated by dividing the number of agreements by the number of agreements plus disagreements. Concordance on occurrences of C-units—independent clauses plus their modifiers—resulted in agreement rates of 98.63% for McDonald's, 99.25% for Late, and 100% for Alien narratives. Concordance on morpheme marking (e.g., alien/*s* = plural vs. alien/*z* = possession) resulted in agreement rates of 99.10% for McDonald's, 91.80% for Late, and 96.42% for Alien narratives. Concordance on the presence of a word (e.g., *hand* vs. *and*) resulted in perceptual agreement rates of 99.85% for McDonald's, 99.45% for Late, and 99.36% for Alien narratives.

Krippendorff's alpha was used to calculate agreement rates for C-unit density, DDM, and EE coding (Hayes & Krippendorff, 2007; Krippendorff, 2004). Krippendorff's alpha accounts for chance agreement as well as the degree of difference between coders and has been used in other studies of spoken narration (Finestack, Palmer, & Abbeduto, 2012; Heilmann, Miller, & Nockerts, 2010; Heilmann, Miller, Nockerts, & Dunaway, 2010). According to Hayes and Krippendorff (2007), Krippendorff's alpha agreement levels greater than .67 are acceptable for tentative conclusions, and those greater than .80 indicate adequate agreement. Thus, adequate agreement levels are higher than acceptable agreement levels.

EE. Examination of each EE element resulted in adequate agreement rates, reaching .96 in McDonald's narratives, .93 in Late narratives, and .97 in Alien narratives. EE subcategories saw agreement rates of .92 for appendages, .90 for orientations, and .95 for evaluations in McDonald's narratives; .81 for appendages, .94 for orientations, and .97 for evaluations in Late narratives; and .89 for appendages, .76 for orientations, and .88 for evaluations in Alien narratives.

DDM. Examination of agreement on presence of morphosyntactic tokens of AAE within each narrative resulted in adequate rates of .94 for McDonald's narratives, .90 for Late narratives, and .83 for Alien narratives.

Statistical analysis. A 2 (grade) × 3 (visual stimulus type) mixed between/within-subjects multivariate ANOVA was conducted. The between-subjects variable was grade level (lower = Grades 2 and 3; upper = Grades 4 and 5). The within-subject variable was visual stimulus type (no visual, picture sequence, single picture). The dependent variables were narrative measures of interest (EE rate, NDW rate, MLU-w, DDM). An alpha of .05 was used to determine the significance of each multivariate test. A Bonferroni-adjusted alpha level of .01 was used to determine the significance of each univariate test.

A correlation analysis was used to explore the relationships among diagnostic risk and overall performance on the four narrative variables: EE rate, NDW rate, MLU-w, and DDM. Scores for the four narrative variables were pooled across visual stimulus type to yield a total mean for each narrative variable. Spearman rank order correlation was used because the DELV-S scores are ordinal, representing three diagnostic risk categories (0 = *low risk,* 1 = *medium risk,* and 2 = *high risk*).

Results

Visual Stimuli Effects on Narrative Language

Performance on the combined narrative variables differed as a function of visual stimulus type, Wilks' lambda = .69, $F(2, 51) = 11.73$, $p < .01$, $\eta_p^2 = .32$. The pairwise comparisons for the main effect of visual stimuli on the combined dependent variables indicated a significant difference between the no-visual condition and both the picture-sequence ($p < .01$) and single-picture ($p < .01$) conditions. There was no main effect of visual stimuli between the picture-sequence and single-picture conditions on the combined dependent variables ($p = .46$).

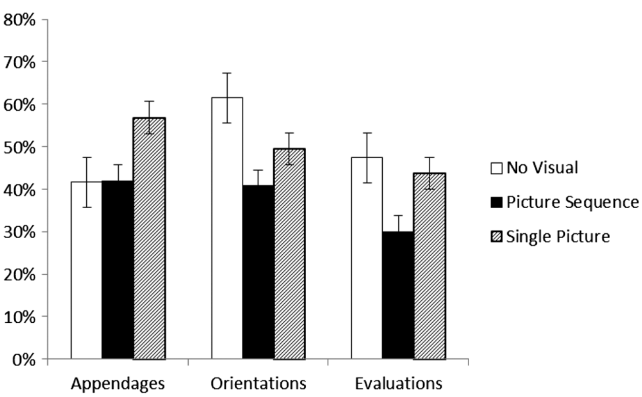
The following univariate tests of visual stimuli effects reached statistical significance: EE rate was higher in the no-visual condition than in either the single-picture or picture-sequence conditions, $F(2, 104) = 7.68$, $p < .016$, $\eta_p^2 = .13$; NDW rate was higher in the no-visual condition than in the picture-sequence condition, which was more lexically diverse than the single-picture condition, $F(2, 104) = 13.85$, $p < .016$, $\eta_p^2 = .21$; and MLU-w was higher in the no-visual condition than in the single-picture or picture-sequence conditions, $F(2, 104) = 9.71$, $p < .016$, $\eta_p^2 = .16$. Effect sizes for these differences were moderate to large. The main effect of visual stimuli on DDM did not reach statistical significance, $F(2, 104) = 4.10$, $p = .019$, $\eta_p^2 = .07$.

Given that visual stimuli effects on EE rate were found, a follow-up repeated-measures ANOVA was conducted. A main effect of visual stimulus was found for two of the three EE subcategories. Rate of orientations was higher in the no-visual condition than in the picture-sequence or single-picture conditions, $F(2, 106) = 18.95$, $p < .016$, $\eta_p^2 = .26$. Rate of evaluations was higher in the no-visual condition than in the picture-sequence or single-picture conditions, $F(2, 106) = 14.29$, $p < .016$, $\eta_p^2 = .21$. Effect sizes for these differences were large. There was no main effect of visual stimulus on appendages, $F(2, 106) = 0.47$, $p = .63$, $\eta_p^2 = .01$. Figure 1 displays the mean percentage of possible points for the three subcategories of EE across the three visual stimulus conditions. Children provided 30% to 62% of possible points for each EE subcategory. Figure 2 displays the mean rate of each EE subcategory across the three visual stimulus conditions.

Grade Effects on Narrative Language

There was a main effect of grade on narrative performance, $F(1, 52) = 11.29$, $p < .01$, $\eta_p^2 = .18$. Univariate tests indicated that NDW rate was higher for children in upper grades, than for children in lower grades, $F(1, 52) = 11.98$,

Figure 1. Mean percentage of the three expressive elaboration subcategories (appendages = 6 points, orientations = 4 points, evaluations = 10 points) across visual conditions.

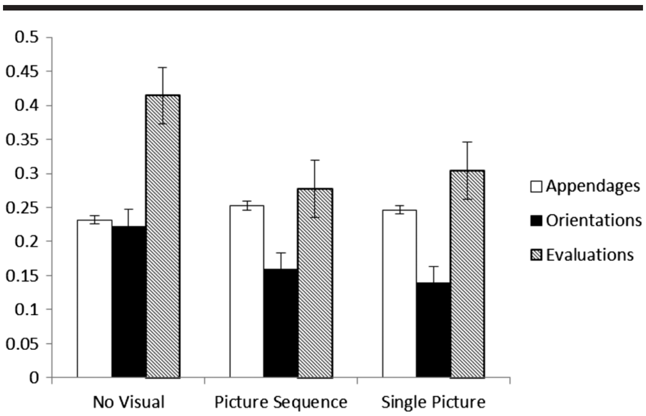


$p < .01$, $\eta_p^2 = .19$. MLU-w was higher for children in upper grades, than for those in lower grades, $F(1, 52) = 11.44$, $p < .01$, $\eta_p^2 = .18$. Effect sizes for these differences were large. There was no main effect of grade for either EE rate, $F(1, 52) = 0.38$, $p = .54$, $\eta_p^2 = .01$, or DDM, $F(1, 52) = 0.44$, $p = .51$, $\eta_p^2 = .01$. Table 2 displays mean performance on the study variables as a function of visual stimulus and grade.

Grade × Visual Stimuli Interaction

The Grade × Visual Stimuli interaction was significant on the combined dependent variables, Wilks's lambda = .70, $F(8, 45) = 2.38$, $p < .05$, $\eta_p^2 = .29$. The interaction between grade and visual stimuli reached statistical significance for DDM, $F(2, 104) = 3.43$, $p < .05$, $\eta_p^2 = .06$. Younger children produced higher rates of dialect density in the no-visual and single-picture conditions than in the picture-sequence condition, but older children did not vary their dialect density

Figure 2. Mean rate of three expressive elaboration subcategories (appendages, orientations, and evaluations) across three visual stimulus conditions.



across the visual contexts. The effect size for this effect was moderate. The interaction between grade and visual stimuli failed to reach statistical significance for the following dependent variables: EE rate, $F(2, 104) = 0.61$, $p = .54$, $\eta_p^2 = .01$; NDW rate, $F(2, 104) = 2.05$, $p = .13$, $\eta_p^2 = .03$; and MLW-w, $F(2, 104) = 1.13$, $p = .33$, $\eta_p^2 = .02$. Figure 3 displays the grade by visual stimuli interaction. Appendix B displays two sample narratives elicited from one younger and one older child in the picture-sequence (Late narrative) condition.

Diagnostic Risk for LI

In addition to the aforementioned analyses, this study examined whether diagnostic risk—as defined by scores on the DELV-S—was related to the narrative variables of interest. There was a negative correlation between diagnostic risk and NDW rate, $\rho = -.46$, $n = 47$, $p < .01$, so that higher levels of diagnostic risk were associated with lower levels of lexical diversity. This correlation was of medium magnitude. Diagnostic risk was not related to the other three narrative variables. Table 3 displays the intercorrelations among diagnostic risk and the narrative variables. Table 4 reports means and standard deviations for narrative variables as a function of diagnostic risk and grade.

Discussion

Similar to other studies of school-age children, this study revealed that African American children with TD produce different rates of narrative macrostructure, microstructure, and dialect density in response to three elicitation conditions: no visual, picture sequence, and single picture. Children produced a higher EE rate in the no-visual condition than in the single-picture or picture-sequence conditions. Children produced higher MLU-w in the no-visual condition than in the single-picture or picture-sequence conditions. Children produced a higher NDW rate in both the no-visual and picture-sequence conditions than in the single-picture condition. A grade by visual stimuli condition interaction was present for DDM so that younger children produced higher rates of AAE in the no-visual and single-picture conditions than in the picture-sequence condition, whereas older children's AAE rates did not differ across the visual conditions. Effect sizes for visual stimuli differences across narrative variables were moderate to large. Last, children with higher diagnostic risk for LI produced lower NDW rates than children with lower diagnostic risk for LI. The following sections discuss each significant finding in turn.

No-Visual Stimulus Yields Better Narrative Structure Than Visual Stimuli

Macrostructure. It was hypothesized that there would be a narrative macrostructure advantage in the no-visual condition on the basis of previous work on narrative writing (Bates, 1991) and spoken narration (Spinillo & Pinto, 1994)

Table 2. Mean (*SD*) narrative performance of school-age African American children across visual stimulus type and grade.

	No visual		Picture sequence		Single picture	
Variable	**Grades 2 and 3**	**Grades 4 and 5**	**Grades 2 and 3**	**Grades 4 and 5**	**Grades 2 and 3**	**Grades 4 and 5**
Macrostructural						
EE rate	.88 (0.21)	.84 (0.18)	.66 (0.26)	.72 (0.23)	.66 (0.25)	.72 (0.34)
Microstructural						
NDW rate	4.64 (0.66)	4.92 (1.04)	3.97 (0.83)	4.93 (1.05)	3.59 (0.88)	4.20 (1.28)
MLU-w	8.28 (1.46)	9.13 (1.55)	6.96 (1.20)	8.46 (1.80)	7.21 (1.50)	7.93 (2.05)
Dialectal						
DDM	.03 (0.02)	.02 (0.01)	.01 (0.01)	.03 (0.02)	.03 (0.03)	.03 (0.02)

Note. No visual = McDonald's story; picture sequence = Late story; single picture = Alien story; EE = expressive elaboration; NDW = number of different word roots; MLU-w = mean length of utterance in words; DDM = dialect density measure.

in school-age children with TD. This hypothesis was supported: EE rate was higher when children told stories in response to no visual than when they told stories in response to either a picture sequence or single picture. Two of the EE subcategories, orientations and evaluations, were also higher when children narrated without a visual cue than when they narrated with either visual stimulus. Children were better able to provide character and setting background and to evaluate the character's motivations and intentions as the character attempts to achieve a goal (e.g., buy food at McDonald's) in the absence of a visual stimulus than in the presence of a picture sequence or single picture. The present study did not find visual stimuli differences on appendages; thus, children were equally able to move the listener from the beginning to the end of their narratives across the three visual conditions.

Several explanations are possible for a narrative macrostructure advantage in the no-visual condition. First, children may have recalled the stories verbatim from the spoken models in the no-visual condition. Perhaps the no-visual condition did not require children to build a story schema; rather, in the no-visual condition, children simply needed to remember and repeat the gist of what was said in the verbal model. This has been referred to as *gist recall* elsewhere (e.g., Schneider & Dubé, 2005). The TNL instructions, which asked children to try to tell the McDonald's story the same way that they heard it in the no-visual condition, may have encouraged gist recall, yielding a better narrative macrostructure than in either of the two visual conditions. A second possibility for the narrative macrostructure advantage in the no-visual context is that children may expend cognitive resources to "read" pictures to extract the gist of the story and then develop the language structuring for the story in the two visual conditions. Indeed, pictured narratives may include varying amounts of macrostructural elements that children must make use of to comprehend story structure, as one new study suggested (Wagner, 2013). Therefore, visual stimuli conditions without prior language structuring may have increased children's cognitive load, reducing their narrative macrostructure performance.

The current study provides exploratory data from African American school-age children that support the work of others on British and Italian children (Spinillo & Pinto, 1994), suggesting that school-age children do not alter their narrative macrostructure performance between single pictures and picture sequences. Taken together, findings from the current study and from Spinillo and Pinto (1994) indicate that single pictures, whether produced by a test manufacturer or a child, are just as supportive of

Figure 3. Grade × Visual Condition interaction on dialect density.

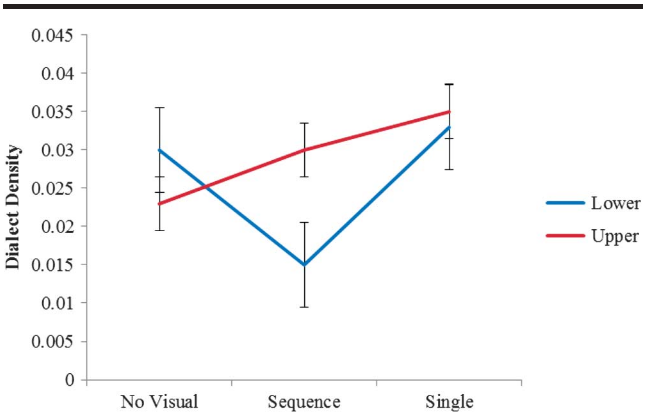


Table 3. Intercorrelations among diagnostic risk and study variables controlling for age.

Variable	1	2	3	4	5
1. Diagnostic risk	—	−.23	−.46*	−.21	.34
2. EE rate		—	.66*	.28	−.03
3. NDW rate			—	.71*	−.11
4. MLU-w				—	−.02
5. DDM					—

Note. Diagnostic risk is based on the Diagnostic Evaluation of Language Variation–Screening Test (0 = *lowest risk,* 1 = *low to medium risk,* 2 = *medium to high risk,* 3 = *highest risk*). EE = expressive elaboration; NDW = number of different word roots; MLU-w = mean length of utterance in words; DDM = dialect density measure.

*$p < .01$.

Table 4. Mean (*SD*) performance on narrative variables across grade and diagnostic risk groups.

Grade	Diagnostic risk	EE rate	NDW rate	MLU-w	DDM
Lower (Grades 2 and 3)	Low	.76 (0.17)	4.32 (0.53)	7.68 (1.00)	.02 (0.01)
	Medium	.72 (0.20)	3.90 (0.44)	7.34 (0.96)	.02 (0.01)
	High	.73 (0.14)	3.89 (0.42)	7.36 (1.00)	.04 (0.02)
Upper (Grades 4 and 5)	Low	.80 (0.11)	5.13 (0.88)	8.76 (1.61)	.03 (0.01)
	Medium	.70 (0.17)	4.18 (0.67)	8.16 (1.64)	.03 (0.01)
	High	.57 (0.04)	3.89 (0.12)	7.69 (0.60)	.02 (0.00)
All		.74 (0.16)	4.23 (0.68)	7.73 (1.18)	.03 (0.02)

Note. Mean scores and standard deviations for each measure are reported. Diagnostic risk is based on the Diagnostic Evaluation of Language Variation–Screening Test. EE = expressive elaboration; NDW = number of different word roots; MLU-w = mean length of utterance in words; DDM = dialect density measure.

narrative macrostructure as a picture sequence. Future studies with other ethnic groups are needed before universal conclusions can be drawn.

Microstructure. Visual stimuli effects were found for both measures of narrative microstructure and are congruent with the hypothesis that narrative microstructure would be higher in the no-visual condition than in the picture-sequence condition, which was based on Nurss and Hough (1985). Indeed, MLU-w followed the same pattern as narrative macrostructure. That is, children produced greater syntactic complexity in the no-visual condition than in either visual condition. Perhaps the sentence structure that they heard in the modeled story was more complex than what they could generate without a verbal model in the two visual contexts. As with narrative macrostructure, children may have engaged in gist recall of syntax in the no-visual condition.

In terms of lexical diversity, children in this study produced a greater NDW rate when narrating in response to no visual cue or to a picture sequence than to a single picture. The findings of the present study are consistent with other studies of narrative lexical diversity in school-age children. For example, Nurss and Hough (1985) found that children produced greater levels of lexical diversity when telling a story without a visual cue than when narrating from either a wordless book or a single picture. Future studies should include a wider variety of visual stimulus types to determine their impact on spoken narrative performance in African American school-age children.

Narrative Dialect Density Differs Across Visual Stimuli in Younger Children

Because visual effects on dialect density have not been directly examined, there was not a firm empirical basis to draw upon for hypothesis development. Yet, it was conjectured that dialect density would be lower in the visual conditions than in the no-visual condition. It was tentatively reasoned that children may interpret the visual stimuli as a literate context (reading and writing) with a high demand for MAE production rather than an oral context (speaking) with a low demand for MAE production.

Our hypothesis was met for younger child narrators but not for older child narrators. That is, a grade by visual stimulus condition interaction on DDM existed so that younger children produced rates of dialect density in the no-visual (M = 0.03, SD = 0.02) and single-picture (M = 0.03, SD = 0.03) conditions that were significantly higher than those produced in the picture-sequence condition (M = 0.01, SD = 0.01). In contrast, African American children in the older group did not dialect shift across visual conditions. As shown in Table 2, these children produced DDM rates that were similar in the no-visual (M = 0.02, SD = 0.01), picture-sequence (M = 0.03, SD = 0.02), and single-picture (M = 0.03, SD = 0.02) conditions.

In the picture-sequence condition, younger African American children produced significantly lower DDM rates than did older children. As illustrated in Appendix B, younger children tended to produce fewer morphosyntactic features of AAE than did older children in the picture-sequence condition. This trend may exist because younger children associated the picture sequence with a book to be read rather than a story to be told. Thus, younger children may have interpreted the picture-sequence condition as a literate context to a greater extent than the single picture, which allows for a less standard narration. Therefore, the results of the current study indicate that, for younger children, dialect density differs across sampling contexts. This finding is consistent with previous research studies of school-age African American children who produced both lower (Ivy & Masterson, 2011) and higher (Thompson et al., 2004) DDM rates than children in the current study, indicating that these children possess the metalinguistic awareness needed to dialect shift and are, accordingly, at promise for academic success.

It is interesting to note that African American children in the current study did not dialect shift on the basis of diagnostic risk for LI because their dialect density rates were unrelated to diagnostic risk. If dialect density is unrelated to risk for LI, then AAE production should not hinder academic performance. However, previous studies suggest that AAE speakers with greater proficiency in MAE tend to perform better on standard literacy tasks than AAE speakers with less proficiency in MAE (Craig, Kolenic, & Hensel, 2014; Craig et al., 2009; N. P. Terry & Connor, 2012). Future research is needed to determine the role that dialect density plays in educational outcomes for school-age

African American children in different regions of the United States.

Diagnostic Risk for LI and Vocabulary Diversity

African American children in the current study were typically developing on the basis of their educational placement in general education classrooms and their performance on standardized tests of cognition, language, and narration. However, standardized measures do not always capture the nuances of language so that linguistic vulnerabilities or strengths go unexposed. The current study revealed that African American children with TD who were at lower risk for LI had greater vocabulary diversity than those at higher risk for LI. This finding extends empirical work suggesting that narrative language—macrostructure and microstructure in particular—distinguishes school-age children with LI from children with TD (Schneider, 1996; Scott & Windsor, 2000). This finding is also consistent with work suggesting that narrative microstructure is more sensitive than macrostructure to language ability in culturally and linguistically diverse children (Burns et al., 2012). Therefore, measures of microstructure, such as NDW rate, should be used alongside standardized measures of language to provide a fuller picture of language ability in African American children.

Limitations

The study potentially confounds narrative task with visual stimuli condition. That is, the present study compares a narrative retell (no-visual condition) with two narrative generations (picture-sequence and single-picture conditions). However, Merritt and Liles (1989) examined narrative macrostructure (story grammar) and microstructure (clause length) differences between the narrative retells and narrative generations of children with and without LI and found that children with TD produced no difference in narrative performance across narrative retell and narrative generation tasks on length-controlled narrative variables. Therefore, it is likely that differences observed in the present study are a function of visual stimuli effects rather than narrative task. Future studies that hold type of narrative (retell, generation) constant across visual stimuli contexts are needed to further substantiate this claim; however, there is evidence suggesting that when school-age children produce narrative retells with and without visual stimuli, performance is improved with verbal, rather than pictorial, support (Schneider & Dubé, 1997, 2005).

Narratives were elicited in the order consistent with standard TNL administration procedures. Therefore, a sequencing effect was present so that the second narrative (Alien) contained more utterances and a greater number of words than the first narrative (Late). To reduce the impact of this sequencing effect, all of the dependent variables of the study were length controlled. Therefore, the visual stimuli type and grade effects on narrative performance were not confounded by narrative length.

Last, expected grade effects were not found for macrostructure. This may be due either to a small sample size with an uneven distribution of participants across grade level or to the measure of macrostructure selected. In the future, other measures of macrostructure should be explored; however, EE has been shown to improve with age in other studies that included a greater number of children (Ukrainetz & Gillam, 2009). Therefore, future studies should examine visual stimuli effects on macrostructure in a larger group of African American children.

Implications

This study has implications for narrative elicitation as well as instruction in vocabulary and code switching (AAE/MAE). Whether in the classroom, clinic, or home environment, caregivers must choose their specific narrative task carefully, depending on the narrative structure they hope to elicit. The current study provides evidence for clinical judgments concerning the identification of children with LI and for decisions about narrative treatment targets through a description of school-age children with TD. The current study suggests that school-age AAE speakers with TD tell stories that include more EE, lexical diversity, and syntax complexity without a visual cue than with a single picture or picture sequence. Of the narrative language variable examined in this study, lexical diversity appears to hold the most diagnostic potential. That is, African American children with TD with greater lexical diversity are at lower risk for LI.

School-based speech-language pathologists (SLPs) may collaborate with classroom teachers to assist children in meeting Common Core benchmarks (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) that require them to tell stories with descriptive detail. Results suggest that African American children produce richer vocabulary either without a visual stimulus or with a picture sequence; therefore, these two contexts should be explored further as supportive of descriptive details. Steele and Mills (2011) provide contextualized activities through which SLPs and classroom teachers can provide direction instruction in vocabulary. For instance, in the context of a book-reading activity in Tier I of the response to intervention (RTI; Butler & Nelson, 2005), educators might provide scaffolds such as student-friendly definitions of words to support gains in vocabulary diversity within children in the general education classroom.

Further, SLPs and teachers may not know the extent to which AAE-speaking children are becoming bidialectal narrators without using different visual stimuli to elicit spoken narratives. The current study indicates that AAE speakers in second and third grades produce a lower rate of dialect density when presented with a picture sequence compared with a single picture or no visual stimulus. Therefore, multiple types of visual conditions should be present to gain a complete picture of narrative performance.

Moreover, results of the current study suggest that visual stimuli consideration may be useful not only in assessing dialect density but also in manipulating dialect density. Given that MAE is preferred in classrooms, efforts should be made to promote competence in bidialectalism so that children have facility with AAE as well as MAE. Although the morphosyntactic rules of AAE and MAE overlap to some extent, they may also differ in ways that tax the working memory of AAE-speaking children (J. M. Terry, Hendrick, Evangelou, & Smith, 2010). It also stands to reason that MAE-speaking teachers and SLPs would also experience an increased cognitive load in processing AAE features that had no MAE counterpart (e.g., invariant *be*). The present study provides preliminary support for the use of picture sequences to support code switching from AAE to MAE in younger school-age children with TD.

Conclusions

Clinicians and researchers should be aware of how visual stimuli may influence the narrative performance of school-age AAE speakers. Narrative macrostructure, microstructure, and dialect density may be attenuated or facilitated as a function of narrating from no visual stimulus, a single picture, or a picture sequence. Thus, narrative elicitation contexts that include both audio and visual tasks are needed to fully describe narrative performance. This research has implications for classroom instruction and language intervention related to narrative performance, which is an important academic benchmark. Findings of the present study suggest that visual stimuli—picture sequences in particular—need to be examined more carefully with respect to their potential utility in supporting different language goals such as improving vocabulary diversity, increasing sentence complexity, and promoting bidialectalism.

Acknowledgments

I am grateful for a Social and Behavioral Sciences Small Grant, which provided seed money to conduct the study. I am also grateful to the children who participated, to the parents who granted them permission, and to the school staff who created space for data collection. Many thanks to the research assistants who helped code narrative data. This study was approved by the Behavioral and Social Sciences Institutional Review Board at The Ohio State University.

References

Adger, C. T., Wolfram, W., & Christian, D. (2007). *Dialects in schools and communities* (2nd ed.). New York, NY: Routledge.

Allen, M. M., Ukrainetz, T. A., & Carswell, A. L. (2012). The narrative language performance of three types of at-risk first-grade readers. *Language, Speech, and Hearing Services in Schools, 43,* 205–221.

Allen, M. S., Kertoy, M. K., Sherblom, J. C., & Pettit, J. M. (1994). Children's narrative productions: A comparison of personal event and fictional stories. *Applied Psycholinguistics, 15,* 149–176.

American Speech-Language-Hearing Association. (2003). *American English dialects* [Technical report]. Available from http://www.asha.org/policy

Bates, L. (1991). The effects on the structure of young children's written narrative of using a sequence of pictures or a single picture as a stimulus. *Reading, 25*(3), 2–10.

Bliss, L. S., Covington, Z., & McCabe, A. (1999). Assessing the narratives of African American children. *Contemporary Issues in Communication Science and Disorders, 26,* 160–167.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2012). *Test of Nonverbal Intelligence* (4th ed.). San Antonio, TX: Pearson.

Burns, F. A., de Villiers, P. A., Pearson, B. Z., & Champion, T. B. (2012). Dialect-neutral indices of narrative cohesion and evaluation. *Language, Speech, and Hearing Services in Schools, 43,* 132–152.

Butler, K., & Nelson, N. (2005). Responsiveness to intervention and the speech-language pathologist [Special issue]. *Topics in Language Disorders, 12*(2). (See six articles on RTI and SLPs.)

Celinska, D. (2009). Narrative voices of early adolescents: Influences of learning disability and cultural background. *International Journal of Special Education, 24,* 150–172.

Champion, T. B. (2003). *Understanding storytelling among African American children: A journey from Africa to America*. Mahwah, NJ: Erlbaum.

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. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research, 47,* 450–463.

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: Erlbaum.

Craig, H. K., Washington, J. A., & Thompson-Porter, C. (1998). Average C-unit lengths in the discourse of African American children from low-income, urban homes. *Journal of Speech, Language, and Hearing Research, 41,* 433–444.

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.

Curenton, S. M. (2006). Oral storytelling: A cultural art that promotes school readiness. *Young Children, 61,* 78–89.

de Villiers, P. (2004). Assessing pragmatic skills in elicited production. *Seminars in Speech and Language, 25,* 57–71.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). San Antonio, TX: Pearson.

Epstein, S., & Phillips, J. (2009). Storytelling skills of children with specific language impairment. *Child Language Teaching and Therapy, 25,* 285–300.

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.

Finestack, L. H., Palmer, M., & Abbeduto, L. (2012). Macrostructural narrative language of adolescents and young adults with Down syndrome or fragile X syndrome. *American Journal of Speech-Language Pathology, 21,* 29–46.

Gazella, J., & Stockman, I. J. (2003). Children's story retelling under different modality and task conditions: Implications for standardizing language sampling procedures. *American Journal of Speech-Language Pathology, 12,* 61–72.

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

Glenn-Applegate, K., Breit-Smith, A., Justice, L. M., & Piasta, S. B. (2010). Artfulness in young children's spoken narratives. *Early Education and Development, 21,* 468–493.

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, and Hearing Services in Schools, 32,* 114–125.

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

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. New York, NY: Cambridge University Press.

Heilmann, J., Miller, J. F., & Nockerts, A. (2010). Sensitivity of narrative organization measures using narrative retells produced by young school-age children. *Language Testing, 27,* 603–626.

Heilmann, J., Miller, J. F., Nockerts, A., & Dunaway, C. (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology, 19,* 154–166.

Hester, E. J. (2010). Narrative correlates of reading comprehension in African American children. *Contemporary Issues in Communication Sciences and Disorders, 37,* 73–85. doi: 1092-5171/10/3701-0073

Hicks, D. (1991). Kinds of narrative: Genre skills among first graders from two communities. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure* (pp. 55–87). Mahwah, NJ: Erlbaum.

Hoffman, L. M., & Gillam, R. B. (2004). Verbal and spatial information processing constraints in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 47,* 114–125.

Hughes, D. L., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.

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.

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology, 15,* 177–191.

Katz, L., & Champion, T. (2008). There's no "1" way to tell a story. In J. C. Scott, D. Y. Straker, & L. Katz (Eds.), *Affirming students' rights to their own language: Bridging language policies and pedagogical practices* (pp. 192–205). New York, NY: Routledge.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research, 30,* 411–433.

Labov, W. (1972). *Language in the inner city*. Philadelphia, PA: University of Pennsylvania Press.

Liles, B. Z., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research, 38,* 415–425.

Loban, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.

Mandler, J., & Johnson, N. (1977). On the remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9,* 111–151.

Masterson, J. J., & Kamhi, A. G. (1991). The effects of sampling conditions on sentence production in normal, reading-disabled, and language-learning-disabled children. *Journal of Speech and Hearing Research, 34,* 549–558.

Merritt, D. D., & Liles, B. Z. (1989). Narrative analysis: Clinical applications of story generation and story retelling. *Journal of Speech and Hearing Disorders, 54,* 438–447.

Michaels, S. (1981). "Sharing time": Children's narrative styles and differential access to literacy. *Language in Society, 10,* 423–442.

Miller, J., & Iglesias, A. (2010). Systematic Analysis of Language Transcripts (Research Version) [Computer software]. Madison, WI: SALT Software.

Mills, M. T. (2015). Narrative performance of gifted African American school-aged children from low-income backgrounds. *American Journal of Speech-Language Pathology, 24,* 36–46.

Mills, M. T., Watkins, R. V., & Washington, J. A. (2013). Structural and dialectal characteristics of the fictional and personal narratives of school-age African American children. *Language, Speech, and Hearing Services in Schools, 44,* 211–223.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards*. Washington, DC: Author.

Nurss, J. R., & Hough, R. A. (1985). Young children's oral language: Effects of task. *Journal of Educational Research, 78,* 280–285.

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York, NY: Plenum Press.

Pratt, M. W., & MacKenzie-Keating, S. (1985). Organizing stories: Effects of development and task difficulty on referential cohesion in narratives. *Developmental Psychology, 21,* 350–356.

Pruitt, S., & Oetting, J. (2009). Past tense marking by African American English speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 52,* 2–15.

Ross, S. H., Oetting, J. B., & Stapleton, B. (2004). Preterite *HAD + V- ED*: A development narratives structure of African American English. *American Speech, 79,* 167–193.

Schachter, R. E., & Craig, H. K. (2013). Students' production of narrative and AAE features during an emergent literacy task. *Language, Speech, and Hearing Services in Schools, 44,* 227–238.

Schneider, P. (1996). Effects of pictures versus orally presented stories on story retellings by children with language impairments. *American Journal of Speech-Language Pathology, 5,* 86–95.

Schneider, P., & Dubé, R. V. (1997). Effect of pictorial versus oral story presentation on children's use of referring expression in retell. *First Language, 17,* 283–302.

Schneider, P., & Dubé, R. V. (2005). Story presentation effects on children's retell content. *American Journal of Speech-Language Pathology, 14,* 52–60.

Scott, C., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research, 43,* 324–339.

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.

Spinillo, A. G., & Pinto, G. (1994). Children's narratives under different conditions: A comparative study. *British Journal of Developmental Psychology, 12,* 177–193.

Steele, S. C., & Mills, M. T. (2011). Vocabulary intervention for school-age children with language impairment: A review of evidence and good practice. *Child Language Teaching & Therapy, 27,* 354–370.

Stein, N., & Glenn, C. (1979). An analysis for story comprehension in elementary school. In R. Freedle (Ed.), *New directions in discourse processing* (Vol. 2, pp. 53–119). Norwood, NJ: Ablex.

Terry, J. M., Hendrick, R., Evangelou, E., & Smith, R. L. (2010). Variable dialect switching among African American children: Inferences about working memory. *Lingua, 120,* 2463–2475.

Terry, N. P., & Connor, C. M. (2012). Changing nonmainstream American English use and early reading achievement from kindergarten to first grade. *American Journal of Speech-Language Pathology, 21,* 78–86.

Terry, N. P., Mills, M. T., Bingham, G. E., Mansour, S., & Marencin, N. (2013). Oral narrative performance of African American prekindergartners who speak nonmainstream American English. *Language, Speech, and Hearing Services in Schools, 44,* 291–305.

Thompson, C. A., Craig, H. K., & Washington, J. A. (2004). Variable production of African American English across oracy and literacy contexts. *Language, Speech, and Hearing Services in Schools, 35,* 269–282.

Ukrainetz, T. A., & Gillam, R. B. (2009). The expressive elaboration of imaginative narratives by children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52,* 883–898.

Ukrainetz, T. A., Justice, L. M., Kaderavek, J. N., Eisenberg, S. L., Gillam, R. B., & Harm, H. M. (2005). The development of expressive elaboration in fictional narratives. *Journal of Speech, Language, and Hearing Research, 48,* 1363–1377.

Wagner, L. (2013). By the numbers: A quantitative content analysis of children's picturebooks. *Frontiers in Psychology, 4,* 1–3. doi:10.3389/fpsyg.2013.00850

Weddington, G. (2010). It's not the language: Alternative explanations of the education gap for African American children. *Topics in Language Disorders, 30,* 48–56.

Wiederholt, J. L., & Bryant, B. R. (1992). *Gray Oral Reading Tests* (3rd ed.). Austin, TX: Pro-Ed.

Appendix A

Expressive Elaboration Scoring

Element	Maximum points	Description	Example
APPENDAGES	6	Signal the beginning, middle, and end of story	
Introducer	1	Indicates beginning of story	*Once, One day*
Abstract	1	Summarizes story before it begins; provides a title for the story	*The Aliens*
Theme	2	Summarizes story within the narrative	*The boy went to sleep last night and then woke up late for school*
Coda	1	Provides a solution to problem in story; provides the effect of an event in the story	*And then he walked to school, But they ran as fast as they can*
Ender	1	A formal conclusion	*The end, That's the end*
ORIENTATIONS	4	Provide information about the characters and where the story takes place (setting)	
Name	1	Refers to characters with specific identifiers	*Jack, Monique, Thomas*
Relation	1	Defines a character's role in terms of relationships or jobs; EXCLUDES possessives (e.g., *his, mom's*)	*Teacher, sister, brother, family of aliens*
External	1	Provides where or when the story takes place; provides the conditions under which story takes places	*At home, school, in the park, at a family reunion picnic, in a town*
Personality	1	Enduring attributes of characters	*The teacher was nice, mean monster, friendly aliens*
EVALUATIONS	10	Make the story events increasingly vivid	
Modifier	2	Adjectives and adverbs EXCLUDING: *some, other, another, one, little, big, good, bad, more, a lot*	*Late, really, weird-looking, dangerous*
Expression	2	Adjectives with literary or idiomatic usage	*They thought they was gonna choke them to death*
Repetition	2	Uses a word or phrase one or more times to add emphasis	*And then it start getting chilly and chillier*
Internal	2	Uses words that reflect temporary intentions, thoughts, emotional motivations, and reactions, EXCLUDING: *tried to, looked*	*Scared, wants, thought, mad*
Dialogue	2	One character speaks directly to another EXCLUDING: *yes* or *no,* indirect speech	*The teacher said, "That's fine because when I was a kid I was late just like you"; And then the girl had grabbed on the little boy and had said, "Come on!"*

Note. From "The Expressive Elaboration of Imaginative Narrative by Children With Specific Language Impairment," by T. A. Ukrainetz and R. B. Gillam, 2009, *Journal of Speech, Language, and Hearing Research, 52,* pp. 883–898. Copyright © 2009 by the American Speech-Language-Hearing Association. Adapted with permission. Appendages (except Theme) and Orientations received up to 1 point. Evaluations and Theme received 1 point for up to two occurrences. Examples are from the narratives in this study sample.

Appendix B

Sample "Late for School" Stories From Lower and Upper Grade Levels

Second-grade child

> Jerry woke up. And he realized that he was late for school. He tried to (pour the pour the) pour his cereal. But he poured too much milk. And when he tied his shoe his string broke. He ran outside. The bus was (already) already left[EW:gone]. So he realized that he had to walk to school. His legs was hurting[SVA]. And he was late for school. (And then his grandma asked if he) And then his mom asked if he (wanted) wanted a ride to school.

Fourth-grade child

> One day this boy had[HAD] woke up. And he (um) noticed that he was late for school. And then he went to go downstairs to go make him[UPC] some cereal, got on some clothes, and tie[PST] his shoe. But his shoestring had[HAD] ripped. And then he grabbed his backpack. And he walked out the door. And (then he was l*) then his bus was already ahead of him. So he walked to school. And his teacher was standing at the door. And she said, "Why are you late?" He said, "Because (um) I had[HAD] went somewhere last night." And she said "OK."

Note. ( ) = verbal maze; * = incomplete word; [ ] = code.

AAE codes:
[AIN]: *ain't*
[PRO]: appositive pronoun
[DON]: completive done
[DMK]: double marking
[MOD]: double copula/auxiliary/modal
[EIT]: existential *it*
[FSB]: *fitna/sposta/bouta/I'ma*
[HAD]: preterite *had*
[ART]: indefinite article
[IBE]: invariant *be*
[NEG]: multiple negation
[REF]: regularized reflexive pronoun
[BEN]: remote past *been*
[SVA]: subject–verb agreement
[UPC]: undifferentiated pronoun case
[ZAR]: zero article
[COP]: zero copula/auxiliary
[ING]: zero *–ing*
[AUX]: zero modal auxiliary
[PST]: zero past tense
[ZPL]: zero plural *–s*
[POS]: zero possessive
[ZPR]: zero preposition
[ZTO]: zero *to*