Exceptional Children
2017, Vol. 84(1) 97–115

DOI: 10.1177/0014402917727248
journals.sagepub.com/home/ecx
SAGE

Elementary Students' Use of Dialect and Reading Achievement: Examining Students With Disabilities

Brandy Gatlin[1] and Jeanne Wanzek[2]

Abstract

Nonmainstream American English, or *dialect*, among children may have important implications for reading research and practice. However, much of the research involving relations between dialect and literacy has analyzed dialect use in only one context and has omitted students with speech, language, and learning disabilities. Consequently, we examined dialect use in an oral narrative and two writing samples in relation to concurrent and longitudinal reading outcomes in a diverse sample of students, including those with diagnosed disabilities. Overall, most students used features of dialect in oral and written language. Dialect use was significantly and negatively predictive of reading outcomes the same year and 2 years later. Moderator analyses indicated a similar relationship between dialect use and reading for students with speech, language, and learning disabilities, suggesting that students with these disabilities who also use dialect may be at increased risk for reading difficulties. Implications for practice and future research are provided.

On the most recent National Assessment of Educational Progress reading test (National Center for Education Statistics, 2015), 82% of African American and 79% of Hispanic fourth-grade students failed to meet proficient levels. Moreover, students from these racial and ethnic minority groups have traditionally been overrepresented in high-incidence disability categories of special education (Donovan & Cross, 2002; U.S. Department of Education, 2014). Despite the fact that African American and Hispanic children are more likely than their non-Hispanic White peers to live in low-income homes (Federal Interagency Forum on Child and Family Statistics, 2013), research has not determined that poverty can solely explain lower literacy performance among many minority students (e.g., Artiles, Kozleski, Trent, Osher, & Ortiz, 2010; Ferguson, 2007). According to the National Research Council's Committee on the Prevention of Reading Difficulties in Young Children (Snow, Burns, & Griffin, 1998), speaking a dialect of English, or nonmainstream American English (NMAE), may be one risk factor for reading difficulties. Many students from racial minority groups speak a dialect of English that differs from mainstream American English (MAE), which is found in most textbooks and is generally used in schools. The purpose of this study was to examine potential relations between oral and written dialect use and reading achievement.

NMAE and Literacy

Investigations into the relationship between NMAE and literacy date back over 40 years (e.g., Goodman & Buck, 1973; Labov, 1969).

[1]University of California, Irvine
[2]Vanderbilt University

Corresponding Author:
Brandy Gatlin, University of California, Irvine, 3200 Education Bldg, Irvine, CA 92697.
E-mail: gatlinb@uci.edu

In particular, one of the more widely recognized forms of NMAE is African American English. Researchers studied African American English extensively after a verdict was reached in what later became known as the Ann Arbor decision (*Martin Luther King Junior Elementary School Children et al. v. Ann Arbor School District*, 1979). The court decision—which ruled in favor of the plaintiffs, citing the school district's failure to consider the linguistic and cultural backgrounds of its students in instructional programming—acted as an impetus to investigations into African American English and literacy with anticipation of potential implications for practice. However, because of weaknesses in studies, including researchers' failure to measure dialect use specifically among children or to determine whether study subjects were indeed speakers of African American English (Troutman & Falk, 1982), past studies that purported to examine a relationship between dialect and literacy among children resulted in inconsistent findings.

Whereas the potential impact of African American English on the development of literacy skills is certainly worth investigating, research suggests that these studies should not be exclusively limited to African American children.

More recent studies have devoted attention to the African American English spoken by children. These studies have aided in establishing the legitimacy of African American English as a language variation (Wolfram, 1998), have helped those in practice to make distinctions between language differences and language disorders or deficits (Seymour, Bland-Stewart, & Green, 1998), and have provided information resulting in a rich corpus of features of the dialect among children (e.g., Craig & Washington, 2004; Oetting & McDonald, 2001). Studies that have examined participant dialect use in various formats have found that many children use dialect with varying frequency in different contexts (Craig, Kolenic, & Hensel, 2014; Washington, Craig, & Kushmaul, 1998). However, most research to date examining the dialect–literacy relationship has focused on dialect use in only one context, specifically oral language, and conclusive evidence regarding associations between the two factors has yet to be attained.

The majority of research on dialect and literacy has focused on African American children. Whereas the potential impact of African American English on the development of literacy skills is certainly worth investigating, research suggests that these studies should not be exclusively limited to African American children. For one, children who are not African American have been shown to use features of African American English, suggesting that the total number of children who use African American English may actually be greater than those who identify as African American (Charity Hudley & Mallinson, 2011). In addition, recent research has converged to demonstrate several similarities between African American English and Southern American English (e.g., Oetting & McDonald, 2002). Perhaps as a result of strong historical ties (Rickford & Rickford, 2000), African American English and Southern American English share several features. Linguists have also noted that some dialect features, especially phonological features, are not necessarily exclusive to speakers of nonmainstream dialects. For instance, a common feature of African American English and Southern American English concerns the lack of pronunciation of the final '*g*' sound in *-ing* words, such as *running* (Wolfram & Schilling-Estes, 2006). However, speakers who use MAE may not fully pronounce the *g* at the end of *-ing* words in their spoken language, especially in rapid or more casual speech. In identifying features of dialect, Wolfram and Schilling-Estes (2006) noted that the mere presence or absence of specific features in a person's language does not identify the person as a dialect speaker; rather, the frequency with which the

person uses that feature is more indicative of his or her dialect classification.

Dialect and Reading

Regarding the effect that dialect may have on literacy acquisition, researchers have suggested that the phonological and morphosyntactic mismatches between NMAE and MAE make the acquisition of literacy skills difficult for speakers of NMAE (Labov, 1995). This theory is supported by research that has found a negative relationship between dialect use and reading scores (e.g., Craig & Washington, 2004). More recently, however, researchers have begun to explore the metalinguistic or linguistic awareness/flexibility theory (e.g., Charity, Scarborough, & Griffin, 2004), which posits that a student's familiarity with MAE and production of MAE in different contexts is related to literacy achievement. In a recent meta-analysis of 18 years of research on dialect and literacy among elementary-aged students, Gatlin and Wanzek (2015) found a negative and moderate association between dialect production and reading skills, indicating an inverse relationship between them. In other words, the more dialect a student tends to use in the production of language, the lower his or her reading scores tend to be. However, as determined by tests for heterogeneity, a significant amount of variation in effect sizes existed among the studies. Gatlin and Wanzek found that the grade-level moderator (primary or intermediate) did not explain variability among the studies. Socioeconomic status also was not a significant moderator in the study, meaning that the percentage of students from lower-income backgrounds within the various studies did not account for differences in effect sizes.

Gatlin and Wanzek (2015) concluded that the findings of the past few years of research on dialect and literacy suggest that dialect variation does appear to relate to reading in some potentially measurable way. Because of the significant amount of heterogeneity in their analysis, coupled with the lack of significance of the study moderators, the authors determined that the relationship may not be a simple one, as suggested by the mismatch theory. Rather, the authors recommended further investigation into the linguistic awareness/flexibility hypothesis, which states that a student's ability or inability to vary dialect use in different contexts is more highly related to literacy performance. As noted by the authors, one potential avenue for examining the linguistic awareness/flexibility hypothesis would be an analysis of students' dialect use in written contexts and an examination of relations to reading achievement. Writing can serve as a mirror into dialect shifting with students who speak NMAE (Washington, 2011). However, only one peer-reviewed study to date has examined elementary students' dialect use in writing as it relates to reading performance. Craig, Zhang, Hensel, and Quinn (2009) found a significant downward shift between dialect use on an oral narrative task and a written narrative task among 165 first- through fifth-grade African American English-speaking students; they also found that students' written dialect use predicted reading achievement, whereas oral dialect use did not. However, Craig and colleagues analyzed dialect use in only one writing sample. In addition, their study, like most research to this point on dialect use, was cross-sectional, examining relations between students' dialect use and literacy measures during the same year. Analyzing dialect within multiple written language samples and examining these relationships with reading achievement in subsequent grades may provide further information on the potential role of dialect shifting in the acquisition and development of literacy skills.

Dialect, Reading, and Students With Disabilities

DeThorne, Petrille, Schatschneider, and Cutting (2010) found that conversational language was a significant predictor of reading achievement scores after controlling for vocabulary, particularly for children with language difficulties. Although DeThorne and colleagues did not include dialect as part of their study, their finding suggests that relationships between spoken language and literacy

could be different for students with language and learning difficulties. In another study, Bishop and Adams (1990) found that children with spoken language difficulties were at increased risk for later reading comprehension difficulties when compared to children without spoken language difficulties. Further, Scarborough (1990) found that children diagnosed with dyslexia had poorer syntactic skills than age-matched controls, as measured by language samples from a naturalistic context. Numerous studies have found that children with language and/or learning disabilities tend to score more poorly on measures of reading skills, including phonological awareness, than peers with typical language development (e.g., Catts, Adlof, Hogan, & Weismer, 2005; DeThorne et al., 2006). Similarly, children with reading difficulties typically score lower on semantic and morphosyntactic measures when compared with children with no documented reading difficulties or disabilities (e.g., Scarborough, 1990).

It is possible that the relationship between reading skills and spoken language, particularly dialect, may differ for students with language and learning difficulties as compared with their peers.

Based on the findings of DeThorne and colleagues (2010), coupled with an overwhelming amount of evidence of the relation between language skills and reading, it is possible that the relationship between reading skills and spoken language, particularly dialect, may differ for students with language and learning difficulties as compared with their peers. This potential difference is certainly worth investigating. However, several studies examining dialect use have intentionally excluded students who were not considered typically developing (e.g., Craig et al., 2009; Terry, Connor, Thomas-Tate, & Love, 2010). In fact, for Gatlin and Wanzek's meta-analysis (2015), disability status was to be used as a study moderator to determine potential differences in relationships for students diagnosed with high-incidence disabilities, such as speech or language impairment or specific learning disability. However, due to the fact that only one published study had included students with disabilities, a moderator analysis was not possible. The paucity of research for students with disabilities regarding the relationship between dialect use and literacy skills highlights this gap in existing research among special populations. Extending the current research to examine associations between dialect and literacy for those who have been identified with a speech and/or language impairment or specific learning disability may provide insight and direction for future research studies, including intervention research. The present study adds to existing literature on the relationship between dialect and reading by examining this relationship for students with language and learning disabilities in comparison to their typically developing peers.

Purpose of Study and Research Questions

The purpose of this study was to analyze elementary students' use of dialect in an oral narrative and written compositions in relation to their concurrent and subsequent reading achievement. Also, we examined these relationships for students with speech and/or language impairments or specific learning disabilities in comparison to their typically developing peers. Specifically, the following research questions guided this study:

Research Question 1: What is the relationship between children's oral and written use of dialect and concurrent and subsequent reading achievement 2 years later?
Research Question 2: Do the preceding relationships differ for students identified with speech and/or language impairments or specific learning disabilities?

Method

Participants

The sample was pulled from students in a mid-sized southeastern city who were participating in a larger National Institute of Child Health and Human Development-funded longitudinal

study on the prevention and identification of learning disabilities (Al Otaiba, Kim, Wanzek, Petscher, & Wagner, 2014). We randomly selected 250 of the 528 students who completed at least one writing assessment to answer the first research question. Three language samples of each randomly selected student were coded for the occurrence of NMAE features. Specifically, we generated random, nonrepeating numbers using Microsoft Excel's random number generator. We assigned each student a number between 1 and 528 and selected students numbered 1 through 250 for this study.

At the first time point, the students were in 69 second- and third-grade classrooms within 11 schools. Similar to the larger group, the average age of the sample was 8.8 years (SD = 0.62). The percentage of students within each demographic category in the randomly selected subsample was proportional to the full sample (χ^2 = 0.08–1.29, p's > .25). Specifically, just over half of the sample (54.8%) were boys. The racial composition of the group was largely African American (62.4%). White students made up 24.8% of the sample, American Indian/Alaska Native 1.6%, and Asian students 0.8%. The remainder of the students (10.4%) identified as multiracial or other; 1.6% was Hispanic any race. The majority of the students were from low-income backgrounds (81.2%), as determined by eligibility for the federal free and reduced-price lunch program.

Just over 15% (n = 38) of the random sample was identified as receiving special education services, including services for speech or language impairment (n = 27, 10.8%) and specific learning disability (n = 6, 2.4%). Using a researcher-developed survey (Speece et al., 2011), classroom teachers rated five of the six students with specific learning disabilities as below or well below average in reading, spelling, and/or writing. In addition, two of these students were dually diagnosed with language impairments. Two students received special education services under the primary category of autism (0.8%; both also received services for language impairments); another three students (1.2%) received special education services, but the category was not specified.

To address the first research question, in addition to analyzing the randomly selected students' reading scores, we coded transcriptions of these students' oral and written language samples to assess the frequency of NMAE dialect use. We then used dialect density in each of the language samples as a predictor in models with reading achievement as the outcome variable. For the second research question, we analyzed the dialect–reading relationship for students within the sample who received special education services for speech/language impairment or specific learning disability who were below average in reading, spelling, and/or writing (n =34) in comparison with the students in the sample with no identified disabilities (n = 213).

Measures

Reading measures. Upon approval from the university and school district Institutional Review Boards, we assessed children during the spring of their second- or third-grade school year initially on reading and language (oral and written) skills and subsequently tested them on the same measures for the next 2 years. We used reading achievement scores from both the first year and the third year in the analyses. Frequency of dialect use from the language samples from the first year were used for the concurrent and longitudinal analyses. To ensure accurate administration and scoring of each test protocol, each tester was required to pass a fidelity check with a senior research assistant and demonstrate 100% accuracy in administration and scoring prior to each round of testing. A second research assistant also double-scored each assessment.

Word identification. To assess word reading ability, the research team administered the Letter-Word Identification subtest of the Woodcock-Johnson Tests of Achievement–Third Edition (WJLW; Woodcock, McGrew, & Mather, 2001), an untimed assessment that requires students to read isolated letters and words of increasing difficulty. The subtest is norm referenced with an established internal consistency reliability of .91 (McGrew, Schrank, & Woodcock, 2007). The second

test, the Test of Word Reading Efficiency–Second Edition (TOWRE; Torgesen, Wagner, & Rashotte, 2012), consists of two subtests, the Sight Word Efficiency and the Phonemic Decoding Efficiency subtest, that form a composite Total Word Reading Efficiency index score. The Sight Word Efficiency subtest measures the number of real words in list form that a student can accurately identify within 45 s. The Phonemic Decoding Efficiency subtest assesses the number of decodable pseudowords that a student can accurately pronounce in 45 s. Alternate-forms reliability for each subtest ranges from .94 to .97 within the 7- to 10-year-old age range.

Comprehension. We assessed reading comprehension using the Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010). For this test, students are instructed to silently read a list of sentences and verify each sentence's accuracy within a time frame of 3 min. The average alternate-forms reliability for the TOSREC is .96 and .88 for second and third grade, respectively. The second test that we used to assess reading comprehension was the Woodcock-Johnson Tests of Achievement–Third Edition Passage Comprehension subtest (WJPC; Woodcock et al., 2001). For this subtest, students are instructed to read short passages of increasing difficulty and to identify missing key words that would make sense in the context of that passage. This subtest is untimed and norm referenced with an established test-retest reliability of .88 (McGrew et al., 2007).

Language measures. To calculate frequency of NMAE use, we analyzed transcriptions of students' language production in an oral narrative task (Task 6 of the Test of Narrative Language; Gillam & Pearson, 2004) and two writing tasks from the first year of testing. For the Test of Narrative Language, which was administered individually, students were shown a picture and instructed to make up their own story using the picture as a prompt. The language samples were audiorecorded and later transcribed with the Systematic Analysis of Language Transcripts (Miller & Iglesias, 2012) software. For the transcriptions, we used standard conventions, which included segmenting each student's spoken language sample into utterances, or communication units (C-units; Loban, 1976).

We asked each student to write two compositions in response to a narrative and an expository writing prompt. These assessments were administered in group format, and students were given 10 min to write on the given prompt. We administered the two prompts on the same day, with the order of administration alternating between groups. For the narrative composition, each student was asked to write a short essay in response to the prompt "One day when I got home from school . . . ". Each test administrator read instructions verbatim to each group, directing students to write a story about something interesting or unusual that happened on a day when they returned home from school.

We administered the writing prompt from the Essay Composition subtest of the Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009) to students to elicit an expository essay. For this subtest, students are instructed to write about their favorite game and to include at least three reasons why they like it. Like the oral narratives, the language samples from the written narrative and written expository were transcribed with Systematic Analysis of Language Transcripts (Miller & Iglesias, 2012). Members of the research team transcribed these writing samples using modified conventions of Systematic Analysis of Language Transcripts, which is mainly used for analyses of oral language production. In particular, instead of using the standard C-unit to segment language samples, we transcribed the data using a method for transcribing utterances in writing, the minimally terminable unit (T-unit; Hunt, 1965), while preserving any errors found within each child's writing.

Dialect density. We coded each language sample for the occurrence of 26 morphosyntactic and 11 phonological NMAE features. This list of features resulted from an extensive review of literature on African American English and Southern American

English and can be found in Appendix A (supplement available online; e.g., Craig & Washington, 2004; Oetting & McDonald, 2001). Two coders—both doctoral students with experience in the fields of education and speech/language pathology and both familiar with features of African American English and Southern American English—worked closely together to establish reliability before coding took place. As suggested by Oetting and McDonald (2002), we calculated dialect density measures (DDMs), a reflection of the proportion of one's overall dialect use in linguistic production (Craig, Washington, & Thomson-Porter, 1998), by dividing the total number of tokens produced by the total number of utterances in the language sample. We used the resulting quotient as the measure of dialect frequency for each language sample. A sample coded transcription of the oral narrative and resulting DDM is included in Appendix A.

We randomly selected 10% of the oral narrative sample to establish initial reliability on the coding of dialectal features, and the number of agreements divided by the sum of agreements and disagreements provided the reliability coefficient. The coders had difficulty establishing initial reliability in the oral language samples, particularly on the phonological features. Upon reviewing previous studies and finding that morphosyntactic dialect features are more commonly coded than phonological ones, we decided to establish reliability separately for morphosyntactic and phonological features. Interobserver reliability for the morphosyntactic features was high and similar to that of other coefficients found in the literature (.95). After three rounds of reliability attempts and subsequent meetings between the two coders, the reliability coefficients ranged from .38 to .77, which meant that the maximum reliability coefficient for the phonological features observed in the oral narratives (.77) was not representative of a desired reliability coefficient of .90 (Nunnally, 1967). However, based on previous research, this reliability was within standards (e.g., Kohler et al., 2007). Discrepancies were discussed, and the two coders agreed on final coding procedures before moving further.

For the written language samples, in addition to morphosyntactic or grammatical features of NMAE (e.g., subject-verb agreement, as in *they was*), we marked spelling patterns that were indicative of phonological NMAE features (e.g., consonant cluster reduction, as in *toed* for *told*). To establish initial reliability, each coder independently rated 10% of the written language samples for the occurrence of the features and coded morphosyntactic and phonological features separately. Initial reliability coefficients of .95 (written narrative) and .99 (written expository) were established on the morphosyntactic features. As with the oral language samples, some difficulty establishing reliability on the phonological features was apparent. Three rounds of reliability were conducted; the reliability coefficients for the coding of the phonological features ranged from .59 to .81 for the narrative and .62 to .88 for the expository essay.

Data Analysis

Because of difficulty establishing reliability on phonological features, we conducted analyses using morphosyntactic DDMs only. We based this decision on previous research that focused on morphosyntactic dialect features only for this age group (e.g., Thomas-Tate & Connor, 2013), citing developmental reasons and the potential confounding of more general spelling errors in writing. Furthermore, as discussed later, several students had phonological dialect features coded in one or both of their written language samples but apparently did not use any dialect features (morphosyntactic or phonological) in their oral language sample. To address the first research question, we conducted analyses using HLM 7.0 software (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) with standard scores from the reading measures—concurrent and subsequent in separate analyses—as dependent variables, using two-level models to account for reading measure scores for students (Level 1) nested within classrooms (Level 2). We considered creating a three-level model with school as a Level 3 variable; however, the intraclass correlation at the school

level was extremely low in unconditional models (ρ's = .0001–.01) when both school and teacher were included as Level 3 and Level 2 variables, respectively. We therefore decided to create Level 2 models using teachers as the nesting variable. Before adding them to the model, DDMs from the language samples were standardized by converting them to *z* scores with a mean of 0 and an *SD* of 1.

To address the second research question, we modified the previous models used to address Research Question 1 by adding a fixed-effect dummy-coded variable of exceptional student education (ESE) eligibility—specifically, eligibility for specific learning disability or speech/language impairment—as a Level 1 moderator to examine potential differences between dialect use and reading achievement scores for students with these identified disabilities. We used the students without any identified disabilities for comparison. We created an interaction term by multiplying the ESE eligibility variable (0 or 1) by the *z*-scored DDM and then tested the interaction to determine if relationships between dialect and reading measures might differ for the ESE students. Because multiple tests were done, analyses were conducted with a Benjamini-Hochberg linear step-up procedure (Benjamini & Hochberg, 1995) to control for Type I error.

Results

Descriptive statistics for the sample are provided in the first two columns of Table 1 for the dialect frequency variable, among the three language samples. A detailed investigation into the data revealed that 205 students (82%) produced at least one feature of dialect during their oral narrative and 190 (76%) produced at least one dialect feature in one or both of their written language samples. A further analysis of the data also revealed that 28 students were found to have phonological dialect features coded in one or both written

Table 1. Descriptive Statistics.

	Larger sample (*n* = 250)		Typically developing students (*n* = 213)		Students with SLI and SLD (*n* = 34)	
Measures	*M* (*SD*)	Min–max	*M* (*SD*)	Min–max	*M* (*SD*)	Min–max
Dialect density measures						
Oral narrative	0.30 (0.29)	0–1.75	0.29 (0.28)	0–1.75	0.37 (0.33)	0–1.20
Written narrative	0.19 (0.22)	0–1.25	0.18 (0.21)	0–1.00	0.25 (0.31)	0–1.25
Written expository	0.12 (0.17)	0–1.00	0.11 (0.15)	0–0.67	0.15 (0.27)	0–1.00
Reading outcomes						
Grades 2 and 3						
WJLW	104.87 (11.74)	62–130	106.55 (10.37)	62–130	94.73 (13.76)	63–119
TOWRE	97.19 (16.01)	53–133	99.31 (14.55)	61–133	85.85 (18.03)	53–121
TOSREC	99.92 (16.59)	55–140	101.74 (15.23)	55–140	88.30 (17.38)	60–123
WJPC	96.24 (10.12)	58–120	97.68 (9.05)	59–120	87.53 (11.80)	58–101
Grades 4 and 5						
WJLW	101.82 (12.16)	63–135	103.30 (11.08)	66–135	90.35 (13.05)	63–109
TOWRE	98.56 (16.01)	54–133	101.02 (14.01)	62–133	80.41 (18.21)	54–113
TOSREC	95.42 (17.00)	55–145	96.54 (16.13)	55–145	83.94 (16.40)	55–111
WJPC	92.09 (10.82)	53–131	93.39 (9.57)	70–131	81.29 (13.29)	53–101

Note. Dialect density measures are features divided by number of utterances. SLI = speech and/or language impairment; SLD = specific learning disabilities; WJLW = Woodcock-Johnson Tests of Achievement–Letter Word Identification; TOWRE = Test of Word Reading Efficiency; TOSREC = Test of Silent Reading Efficiency and Comprehension; WJPC = Woodcock-Johnson Tests of Achievement–Passage Comprehension.

language samples but no features coded in their oral language sample. In particular, these features were spelling errors that were mostly coded as the final consonant absence or consonant cluster reduction. As indicated by the relatively large standard deviations in comparison with means, students produced a range of frequencies in their dialect use in all three language samples. Pearson's product-moment correlations among the language samples ranged from nonsignificant ($r = .10$, $p = .11$) to small (r's = .14–.23, p's < .05; see Table 2, lower diagonal).

Preliminary analyses also revealed that 97% of African American students used dialect in at least one language sample and 94% of students who qualified for free or reduced-price lunch used at least one instance of dialect in their language samples. As far as frequency of dialect use in the language samples, in general, African American students used a significantly greater amount of dialect in all three language samples (γ's = .10–.41, p's < .01) than students who were not African American. Students from lower income backgrounds, as indicated by eligibility for the free and reduced-price lunch program, tended to use more dialect on the oral narrative ($\gamma = .28$, $p < .001$) and the written narrative ($\gamma = .15$, $p < .01$) than students who were not from lower-income homes. Differences were not significant between these students on the written expository language sample ($\gamma = .06$, $p = .14$).

Mean student scores on each of the Year 1 reading measures generally fell within the average range, and correlations among the reading measures were large in magnitude (r's = .72–.80, p's < .001). The correlations between the frequency of dialect in the language samples and the reading measures indicated that dialect use was significantly and negatively related to each reading measure with a small to moderate magnitude range (r's = –.12 to –.36, p's < .05).

Dialect Density and Concurrent Reading Achievement

Research Question 1 involved two-level HLM analyses examining the relative contributions of dialect use in all three language samples to performance on the reading measures. Overall, the unconditional models (Table 3, Model 1) demonstrated weighted scores for each reading measure within the average range, with 24% to 32% of variance among scores at the classroom level, which was significant for each measure (τ's = 29.32–81.52, p's < .05). As a preliminary step, before all DDM predictors were added to the models simultaneously, we examined the direct effects of dialect use in each language sample separately. In examining the absolute contributions of dialect density in each language task, we found that dialect density was a significant and negative predictor on all reading assessments. In other words, for every unit increase in dialect density on the language samples, scores were generally expected to decrease on each reading variable. For the WJLW, DDMs from all three language samples were significant: oral narrative ($\gamma = -3.59$, $p < .001$), written narrative ($\gamma = -3.09$, $p < .001$), and written expository ($\gamma = -2.56$, $p < .001$). DDMs from two language samples were significant predictors for performance on the TOWRE: the oral narrative ($\gamma = -3.72$, $p < .001$) and the written narrative ($\gamma = -3.48$, $p < .001$). DDMs from the written favorite-game expository were not significant for this measure ($\gamma = -1.25$, $p = .19$). Similarly, DDMs from the oral narrative ($\gamma = -3.77$, $p < .001$) and the written narrative ($\gamma = -2.67$, $p < .01$) were significant predictors for performance on the TOSREC, but DDM from the written expository was not significant ($\gamma = -1.93$, $p = .05$). On the WJPC, all three language sample DDMs were significant: oral narrative ($\gamma = -3.19$, $p < .001$), written narrative ($\gamma = -2.37$, $p < .001$), and written expository ($\gamma = -2.08$, $p < .001$). Pseudo-R^2 values—or the amount of additional variance at the individual level explained over the previous model—ranged from 0 to .11, indicating 0% to 11% variance accounted for by adding the predictors to the model.

The next step was to examine the relative contributions of dialect density in the language samples within the context of one another to performance on the reading measures. As shown in Table 3 (Model 2), when

Table 2. Correlation Matrix for Dependent and Independent Study Variables.

Measures	1	2	3	4	5	6	7	8	9	10	11
Dialect density measures											
1. Oral narrative		.06	.02	**−.51**	**−.45**	−.23	**−.42**	**−.53**	−.27	−.11	−.34
2. Written narrative	**.23**		−.09	−.36	−.29	−.26	−.34	−.11	.07	.07	.09
3. Written expository	.10	**.14**		−.25	−.15	−.09	−.28	**−.60**	−.28	−.06	**−.55**
Reading: Grades 2 and 3											
4. WJLW	**−.34**	**−.31**	**−.24**		**.88**	**.69**	**.86**	**.93**	**.75**	**.53**	**.65**
5. TOWRE	**−.28**	**−.31**	**−.12**	**.80**		**.78**	**.81**	**.88**	**.95**	**.73**	**.73**
6. TOSREC	**−.28**	**−.25**	**−.16**	**.70**	**.75**		**.72**	**.73**	**.76**	**.86**	**.70**
7. WJPC	**−.36**	**−.31**	**−.23**	**.75**	**.73**	**.72**		**.88**	**.81**	**.67**	**.87**
Reading: Grades 4 and 5											
8. WJLW	**−.36**	**−.33**	**−.28**	**.80**	**.77**	**.74**	**.79**		**.75**	**.58**	**.81**
9. TOWRE	**−.19**	**−.27**	**−.14**	**.70**	**.87**	**.72**	**.72**	**.71**		**.69**	**.74**
10. TOSREC	**−.32**	**−.26**	**−.18**	**.56**	**.66**	**.81**	**.67**	**.72**	**.66**		**.70**
11. WJPC	**−.42**	**−.24**	**−.32**	**.65**	**.61**	**.68**	**.72**	**.75**	**.57**	**.65**	

Note. Lower diagonal contains correlation coefficients for full sample (n = 250); upper diagonal contains correlation coefficients for students with speech/language and learning disabilities (n = 34). Correlation coefficients for typically developing sample (n = 213) were nearly identical to that of the full sample. Bolded correlation coefficients are significant at $p < .05$ level. Dialect density measures are features divided by number of utterances. WJLW = Woodcock-Johnson Tests of Achievement–Letter Word Identification; TOWRE = Test of Word Reading Efficiency; TOSREC = Test of Silent Reading Efficiency and Comprehension; WJPC = Woodcock-Johnson Tests of Achievement–Passage Comprehension.

Table 3. Hierarchical Linear Modeling for Dialect Density Measures and Reading Achievement Measures.

	WJLW		TOWRE		TOSREC		WJPC	
Model Specifications	CONC	LONG	CONC	LONG	CONC	LONG	CONC	LONG
Model 1								
Intercept	**104.26**	**101.34**	**96.22**	**98.14**	**99.21**	**99.01**	**95.78**	**91.68**
Variance								
Level 1	105.19	109.10	172.95	198.87	192.34	226.51	73.07	98.38
Level 2	**33.15**	**37.93**	**81.52**	**56.34**	**77.83**	**62.02**	**29.32**	**17.32**
Model 2								
Intercept	**105.32**	**101.93**	**97.43**	**99.18**	**100.16**	**95.54**	**96.49**	**92.37**
Oral narrative DDM	**−2.89**	**−3.21**	**−3.19**	−2.14	**−3.62**	**−5.10**	**−2.84**	**−4.00**
Written narrative DDM	**−2.18**	**−2.52**	**−2.87**	**−3.66**	−2.03	−3.08	**−1.62**	−1.02
Written expository DDM	−1.74	−1.68	—	—	—	—	−1.36	−1.65
Variance								
Level 1	93.16	87.56	160.12	178.37	180.09	192.96	64.37	73.43
Level 2	13.68	24.54	**48.89**	31.97	**46.80**	56.30	12.84	9.13
Model 3								
Intercept	**106.51**	**103.02**	**98.83**	**100.64**	**101.28**	**95.95**	**97.38**	**93.27**
ESE eligibility	**−7.92**	**−10.86**	−6.10	**−15.08**	**−6.87**	−8.95	**−6.25**	**−9.47**
Oral narrative DDM	**−1.94**	**−2.61**	−2.11	−1.29	**−3.04**	**−5.09**	**−2.30**	**−3.89**
Written narrative DDM	**−2.05**	**−2.51**	**−2.75**	**−4.21**	−1.70	**−3.90**	**−1.63**	−1.15
Written expository DDM	−1.65	−1.30	—	—	—	—	−1.07	−1.46
ESE × Oral Narrative DDM	**−4.36**	−3.09	**−6.23**	−5.11	−2.62	1.41	−2.55	0.39
ESE × Written Narrative DDM	−1.70	1.72	−0.13	6.19	0.61	5.89	−0.91	1.99
ESE × Written Expository DDM	0.06	−2.89	—	—	—	—	−0.45	−0.94
Variance								
Level 1	89.16	72.78	158.39	158.05	180.23	188.57	66.51	52.96
Level 2	6.32	27.00	**32.47**	22.95	**37.27**	50.89	7.74	5.43

Note. Bolded coefficients are significant upon Benjamini-Hochberg correction for Type I error. WJLW = Woodcock-Johnson Tests of Achievement–Letter Word Identification; TOWRE = Test of Word Reading Efficiency; TOSREC = Test of Silent Reading Efficiency and Comprehension; WJPC = Woodcock-Johnson Tests of Achievement–Passage Comprehension; CONC = concurrent; LONG = longitudinal; DDM = dialect density measure (features divided by number of utterances); ESE = exceptional student education.

all three measures were included simultaneously in the model, DDMs from two of the language samples were significant predictors for the WJLW: the oral narrative and the written narrative (DDMs from the written expository were not significant). Because DDMs from the written expository were not significant for the TOWRE or the TOSREC in the individual analyses described earlier, we included only students' oral and written narrative DDMs in the simultaneous model. Both predictors remained significant in the model examining their relative contributions to the score on the TOWRE. Oral narrative DDM

significantly predicted a decrease in scores on the TOSREC, but written narrative DDM did not. The degree of dialect use in the oral narrative and the written narrative were significant predictors for the WJPC; written expository dialect use was not significant. For each reading measure, the conditional models were a significant improvement over the unconditional models, $\Delta\chi^2(3) = 117.50–158.02$ (p's < .001), and each explained 6% to 12% of variance at the student level and 39.9% to 58.7% variance at the classroom level. Classroom-level variance was significant for the TOWRE and the TOSREC.

Dialect Density and Longitudinal Reading Achievement

Of the original 250 students from the random sample, 153 were tested 2 years later during their fourth- or fifth-grade year. These students were spread across 63 classrooms in 17 schools. According to chi-square tests for observed versus expected frequencies, this sample was proportional to the previous sample as far as demographic variables were concerned ($\chi^2 = 0.01–4.16$, p's > .25). Means and standard deviations for these reading scores are also displayed in Table 1. Like the scores from 2 years before, scores fell within the average range for the sample. However, on the WJPC, scores were generally in the low-average range. Correlation coefficients (Table 2) for the language sample DDMs and the fourth- and fifth-grade reading scores were found to be similar to those in the previous analysis among DDMs and the concurrent reading scores, ranging from nonsignificant (r = .14, p = .07) to small to moderate (r's = –.18 to –.42, p's < .05).

To examine the relationship between dialect density and longitudinal reading, we repeated the models that simultaneously analyzed dialect density in each language task in relation to second- and third-grade reading achievement, this time using scores from students' fourth- or fifth-grade year. In the models examining the relative contribution of dialect density from the language samples, the longitudinal results were similar to those found in concurrent reading measures analyses (Table 3). On the WJLW, DDMs from the oral narrative and the written narrative were significant predictors. Written expository dialect density was not significant. Dialect density in the oral narrative was not significant for the TOWRE, but written narrative DDMs were significant. In contrast, dialect density on the oral narrative was significant for the TOSREC, whereas written narrative DDMs were not significant. Finally, DDMs from the oral narrative were significant for the WJPC, but written narrative DDMs were not. As in the previous model analyzing reading performance during the second- and third-grade years, written expository dialect density was not a significant predictor for the WJPC.

Students With Speech and/or Language Impairments or Specific Learning Disabilities

To address the second research question, we first calculated the means and standard deviations with correlations for each dialect density variable for the sample of students with speech, language, and learning disabilities (Table 1 and upper diagonal of Table 2). As shown, dialect use among the students with disabilities was slightly more frequent than that of the larger sample. Similar to the larger sample, a substantial amount of variation in students' dialect use within each language sample was present. Scores on the reading measures were in the range of low average to average, generally lower than those of the larger sample. Correlations between dialect density in each language sample and reading assessments for students with speech/language or learning disabilities ranged from marginal to not significant (r's = .07 to –.36, p's > .07) to small to moderate (r's = –.42 to .60, p's < .03).

Next, we conducted multilevel analyses to examine whether the relationship between dialect frequency and reading achievement differed for students with speech and/or language impairments or specific learning disabilities. We repeated the models including dialect density from all three language samples (two for

the TOWRE and TOSREC) for Research Questions 1 and 2, but this time we included a dummy-coded variable for speech/language impairment or specific learning disability and interaction terms between ESE eligibility and DDMs from each language sample. As shown in Table 3 (Model 3), for the WJLW, identification as ESE (speech/language or learning disability) was a significant predictor of decrease in scores, as well as oral narrative and written narrative dialect use. However, dialect frequency in the written expository samples was not significant. The interaction term between speech/language or learning disability and dialect density was significant for the oral narrative but not for the written language samples, indicating that the relationship between oral dialect use and reading was different for the ESE students but not between written dialect use and reading. When we analyzed longitudinal reading outcome models for the WJLW, identification as speech/language impaired or learning disabled and oral narrative and written narrative dialect density were significant predictors for scores. Dialect use in the written expository was not significant. Interaction terms were not significant.

In the model that analyzed dialect use from the oral and written narrative language samples in relation to the TOWRE, ESE eligibility was not a significant predictor for reading scores, indicating that student identification as speech/language or learning disabled did not predict performance on this measure. Oral dialect was also not significant, but written narrative DDM was significant. The interaction term between ESE and oral narrative DDM was significant, but the interaction term between ESE and written narrative DDM was not significant. For longitudinal achievement scores, identification as ESE was significant, and DDMs from the written narrative predicted performance on the TOWRE. Oral narrative DDMs were not significant, and neither were the interaction terms.

When DDMs from the oral and written narrative language samples were analyzed simultaneously for the TOSREC, ESE eligibility was a significant predictor. The oral narrative DDM variable was significant, but the written narrative DDM was not significant. Interaction terms were also not significant. For performance on the TOSREC when students were in Grades 4 and 5, ESE eligibility was not significant. However, oral dialect density and dialect use in the written narrative were significant predictor variables. Interaction terms were not significant.

For the WJPC, speech/language impairment or learning disability was a significant predictor of decrease in scores on WJPC, as were DDMs from the oral and the written narrative. The written expository was not significant. Interaction terms were not significant either. For the models involving scores on the WJPC 2 years later, when all three dialect variables were included in the same model with ESE eligibility, ESE eligibility was significant, as was dialect use in the oral narrative. Written narrative DDMs and written expository DDMs were not significant. The interaction terms were not significant.

Discussion

Based on the potentially important relationship between dialect and literacy, the purpose of this study was to analyze dialect use among a sample of middle elementary grade students and to examine potential relationships between dialect use and scores on reading achievement measures. Specifically, we examined relationships between students' dialect use in three language samples and four measures of reading achievement, during the same school year and 2 years later. We found that dialect density was generally negatively related to performance on the reading measures included in this study. The present study adds to previous research in the area by examining longitudinal reading outcomes and including students who were more diverse than most previous studies, which largely consisted of only African American students. Second, this study investigated potential differences in relationships between dialect and reading for students who had been diagnosed with a speech and/or language impairment or a specific learning disability. Findings showed that relationships between dialect use and reading scores were different for students with

these high-incidence disabilities on two of the four measures included in the study.

As determined by means and standard deviations of the DDM variables, several students used dialect in their spoken and written language. Correlations between dialect frequency variables and the reading measures were all significant and negative, ranging from small to medium. This finding was not surprising given that the effect size found in the meta-analysis of previous studies analyzing dialect use and reading outcome was moderate and negative (Gatlin & Wanzek, 2015). In addition, HLM analyses revealed that a standard deviation increase in dialect use predicted a decrease in reading performance ranging from roughly 2 to 5 standard score points. Results were similar for the analyses involving dialect use and reading achievement concurrently and for the same students 2 years later. Although explanation regarding the nature of the relationship between dialect and reading is premature, this finding is important in that it confirms a significant relationship between the two factors that is sustained over time.

The finding of a negative relationship between dialect density and reading scores is in line with the linguistic interference theory (Labov, 1995) and consistent with past studies (e.g., Charity et al., 2004). However, in this study, in the more sophisticated multilevel models, there remained a substantial amount of variability, even with dialect predictors from all three language samples included. This finding, coupled with the results from the previous meta-analysis (Gatlin & Wanzek, 2015) that indicated a significant amount of heterogeneity among studies, implies that it would be difficult to conclude that the relationship between dialect use and reading achievement can be explained simply by differences, or mismatches, in speech and print. Instead, it appears that a more complex relationship may exist and that the findings may be more in line with the linguistic awareness/flexibility hypothesis (Terry & Scarborough, 2011). This theory would be supported by more complex relations between dialect use and literacy scores, such as varying correlations between dialect in different contexts and literacy performance, as found in this study. In addition, slope parameters were different for each of the language sample's dialect variables on reading measures when controlling for one another, which would indicate that, in the presence of one another, dialect from the various language tasks has a different relationship with reading scores, concurrently and longitudinally 2 years later. This finding is important in that it confirms the need for future studies to examine dialect in various contexts as it relates to reading.

Although we found that dialect use was negatively related to scores of reading achievement, the findings—namely, the predictive nature of dialect in the oral and written narratives—are in contrast to the results of a previous study conducted by Craig and colleagues (2009). In this study, the researchers found that by controlling for dialect in writing, dialect use in an oral language task was not significantly related to reading scores. Yet, students' frequency of dialect use in the written language task did predict a decrease in reading scores. In their study, however, Craig and colleagues' DDMs included total dialect density, a composite of morphosyntactic, phonological, and what are referred to as combination morphosyntactic and phonological African American English features (e.g., the zero-past tense).

Our reasons for analyzing only morphosyntactic features included the lower-than-desirable reliability coefficient in the coding of phonological features and the potential confounding of students' general spelling skills for the age of the sample. Craig and colleagues reported high reliability (91%–100%) for the dialect features coded in their study but did not report separate reliability coefficients for morphosyntactic, phonological, or combination features. Because the authors did not control for spelling skills, there is a possibility that general spelling ability may have been confounded with phonological dialect use in their students' writing samples, thus helping to explain the significant relationship of reading with written dialect as opposed to oral dialect use. In addition, Craig and colleagues controlled for socioeconomic status, general

oral language abilities, and writing skills and included only African American English speakers in their study, which might explain differences in findings between the studies.

For students receiving services for speech/language or learning impairments, the association between oral dialect use and performance on measures of word identification may be even greater than that of their typically developing peers.

In almost all models, students with speech/language and learning impairments had significantly lower scores than their peers who were not eligible for ESE services. Dialect use in the oral and/or written narrative remained significant in all models including the ESE variables. In addition, in the concurrent models, interactions between ESE eligibility and dialect density in the oral narrative and the two measures of word identification (WJLW and TOWRE) were significant. This finding indicates that for students receiving services for speech/language or learning impairments, the association between oral dialect use and performance on measures of word identification may be even greater than that of their typically developing peers. Interaction terms between written dialect use and ESE eligibility were not significant for the two measures of word identification, suggesting that the relationship between written dialect use and word identification did not differ for students identified with these particular disabilities. Interaction terms were also not significant for the two measures of reading comprehension (TOSREC and WJPC), which indicates that dialect density appears to have a similar negative relationship with reading comprehension for students with these disabilities as it does for typically developing students. Together, these findings suggest that students who are eligible for services under the categories of speech, language, or learning impairment who also use significant amounts of dialect in their speech and/or writing may be at even further risk of low reading scores than their peers who receive services under these high-incidence categories who do not use dialect. In addition, there may be some measurable difference specifically for students with speech/language impairments or learning disabilities who speak with a dialect on performance on word identification measures, a finding that warrants further investigation.

Implications for Practice and Future Research

According to the linguistic awareness/flexibility theory, dialect use is not a problem in and of itself; however, it may serve as a marker or evidence of potential reading difficulties, which has important implications for practice and research. Researchers have suggested that students who are dialect users may benefit if they become bidialectal—that is, able to navigate between nonstandard dialects and more standard dialects in various settings (e.g., Craig et al., 2009; Terry et al., 2010). Being bidialectal would have similarities to bilingualism, as instruction in standard English would ideally lead to proficiency in using standard English when appropriate or expected, without devaluing the student's home or heritage language. It will be important for educators and researchers to be mindful of the sensitive nature of the topic of language variation (see Wolfram, 1998) and to consider, and also respect, students' linguistic and cultural backgrounds in designing future research studies and planning instruction. Charity Hudley and Mallinson (2011) suggested that when addressing issues of language varieties in classrooms, educators need to talk about different patterns of English as opposed to correct versus incorrect English. Delpit (2006) and other researchers have suggested using writing, as opposed to spoken language, as an arena to address the teaching of standard English forms. Because writing lends itself to editing, unlike unplanned spoken language, dialogue on differences between language varieties and mainstream English may be more suitable during writing instruction.

The linguistic awareness/flexibility hypothesis provides implications for the malleability

of dialect use in different contexts. Terry, Connor, Petscher, and Conlin (2012) suggested that instruction geared toward improving literacy among young children who use dialect need not point out dialect differences specifically but rather should encourage children to become more attuned to language itself, thereby increasing metalinguistic awareness among the students. Currently, few strategies that focus on dialect awareness have been empirically tested among elementary students. Recently, Connor and colleagues (2014) employed an intervention study designed to increase dialect awareness in their written language among second- through fourth-grade students. The researchers found that students who received the dialect awareness intervention used fewer features of dialect in their posttest writing samples than those students who did not receive the intervention. However, the researchers did not examine whether increased awareness of differences in dialect features and standard English transferred to improved reading outcomes. Experimental intervention studies implementing these strategies or a combination of these strategies and applying various reading outcomes would provide information regarding what methods might be more effective and for whom.

The relationship between dialect and reading may be better conceptualized as reading performance being predictive of dialect use. To our knowledge, no study has investigated literacy-related predictors with dialect use as a dependent variable. If dialect use is a marker for reading difficulties, as suggested by the metalinguistic awareness/flexibility theory, then future studies might investigate predictors of the frequency of dialect use in various contexts. Finally, longitudinal research examining dialect use and variability over time in written language samples may not only provide information on developmental growth trajectories of dialect use but also have the potential for examining growth patterns in dialect use in relation to reading development.

Limitations

Although informative, this study was not without limitations. For one, the study was correlational; therefore, the results do not imply causal relationships between dialect use and outcomes on measures of reading achievement. However, the correlational findings of the study could provide implications for possible future studies, including intervention research, which has not taken place widely within this area of reading research. We acknowledge as a limitation the fact that reliability was difficult to establish, especially among the phonological features. We controlled for this limitation, however, by conducting analyses with morphosyntactic features, which are more explicit and obvious. It will be important for researchers who engage in work directed toward analyzing dialect use to be mindful of these potential difficulties. Providing extended training or buffering scheduling to accommodate for repeated practice in coding phonological features may help improve reliability. Finally, the sample of students with disabilities included only those with speech/language or learning disabilities and may have been too small to detect significant effects, particularly significant interaction effects. Future research may obtain a larger sample size and include students with other disabilities that may affect reading acquisition.

Conclusion

This study extends the current research base on dialect and reading in several important ways. First, most studies investigating relations between dialect use and reading have involved children's spoken use of dialect. In this study, we found evidence that dialect use in an oral and a written narrative was predictive of reading performance. In addition, the longitudinal portion of the study adds to the existing literature on dialect and reading investigations, which have largely consisted of cross-sectional design studies. Finally, the study found that for students with speech, language, or learning disabilities, the impact of dialect use may be greater on measures of word identification than for students with no diagnosed disabilities. However, the relationship is generally the same as it is for students with no diagnosed disabilities on measures of reading comprehension. Although this study showed that students use

dialect differently in various contexts and that there is a significant and negative relationship between the two factors, other variables apparently need to be simultaneously analyzed to help explain some of the variability present. Furthermore, the direction of the relationship between dialect and reading is not certain. Although the field has made considerable progress since the influential 1979 Ann Arbor decision, the results of the meta-analysis and the present study suggest that there is still more to learn regarding language variation among children and its relations to reading acquisition to be able to successfully address the achievement gap.

References

Al Otaiba, S., Kim, Y.-S., Wanzek, J., Petscher, Y., & Wagner, R. K. (2014). Long term effects of first grade multi-tier intervention. *Journal of Research on Educational Effectiveness*, *7*, 250–267. doi:10.1080/19345747.2014.906692

Artiles, A. J., Kozleski, E. B., Trent, S. C., Osher, D., & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968–2008: A critique of underlying views of culture. *Exceptional Children*, *76*, 279–299. doi:10.1177/001440291007600303

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*, 289–300.

Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *31*, 1027–1050. doi:10.1111/j.1469-7610.1990.tb00844.x

Catts, H. W., Adlof, S. M, Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research*, *48*, 1378–1396. doi:10.1044/1092-4388(2005/096)

Charity, A. H., Scarborough, H. S., & Griffin, D. M. (2004). Familiarity with school English in African American children and its relation to early reading achievement. *Child Development*, *75*, 1340–1356. doi:10.1111/j.1467-8624.2004.00744.x

Charity Hudley, A. H., & Mallinson, C. (2011). *Understanding English language variation in U.S. schools*. New York, NY: Teachers College Press.

Connor, C. M., Phillips, B. M., Kaschak, M., Apel, K., Kim, Y.-S., Al Otaiba, S., . . . Lonigan, C. J. (2014). Comprehension tools for teachers: Reading for understanding from prekindergarten through fourth grade. *Educational Psychology Review*, *26*, 379–401. doi:10.1007/s10648-014-9267-1

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. doi:10.1044/1092-4388(2013/12-0157)

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. doi:10.1044/1092-4388(2004/036)

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*, *43*, 366–379. doi:10.1044/jslhr.4102.433

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. doi:10.1044/1092-4388(2009/08-0056)

Delpit, L. D. (2006). *Other people's children: Cultural conflict in the classroom*. New York, NY: New Press.

Dethorne, L. S., Hart, S. A., Petrill, S. A., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., & Davison, M. D. (2006). Children's history of speech-language difficulties: Genetic influences and associations with reading-related measures. *Journal of Speech, Language, and Hearing Research, 49,* 1280–1293. doi:10.1044/1092-4388(2006/092)

Dethorne, L. S., Petrill, S. A., Schatschneider, C., & Cutting, L. (2010). Conversational language use as a predictor of reading development: Language history as a moderating variable. *Journal of Speech, Language, and Hearing Research, 53,* 209–223. doi:10.1044/1092-4388(2009/08-0060)

Donovan, S., & Cross, C. (2002). *Minority students in special and gifted education*. Washington, DC: National Academies Press.

Federal Interagency Forum on Child and Family Statistics. (2013). *America's children: Key national indicators of well-being, 2013*. Washington, DC: U.S. Government Printing Office. Retrieved from https://www.childstats.gov/pdf/ac2013/ac_13.pdf

Ferguson, R. F. (2007). *Toward excellence with equity: An emerging vision for closing the achievement gap*. Boston, MA: Harvard Education Press.

Gatlin, B., & Wanzek, J. (2015). Relations among children's use of dialect and literacy skills: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *58*, 1306–1318. doi:10.1044/2015_JSLHR-L-14-0311

Gillam, R. B., & Pearson, N. A. (2004). *Test of narrative language*. Austin, TX: PRO-ED.

Goodman, K. S., & Buck, C. (1973). Dialect barriers to reading comprehension revisited. *The Reading Teacher*, *27*(1), 6–12.

Hunt, K. (1965). *Grammatical structures written at three grade levels*. Champaign, IL: National Council of Teachers of English.

Kohler, C. T., Bahr, R. H., Silliman, E. R., Bryant, J. B., Apel, K., & Wilkinson, L. C. (2007). African American English dialect and performance on nonword spelling and phonemic awareness tasks. *American Journal of Speech-Language Pathology*, *16*, 157–168. doi:10.1044/1058-0360(2007/020)

Labov, W. (1969). *The study of nonstandard English*. Washington, DC: National Council of Teachers of English.

Labov, W. (1995). Can reading failure be reversed? A linguistic approach to the question. In V. Gadsden, & D. Wagner (Eds.), *Literacy among African American youth* (pp. 39–68). Cresskill, NJ: Hampton Press.

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

Martin Luther King Junior Elementary School Children et al. v. Ann Arbor School District, 473 F. Supp. 1371 (1979).

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Technical manual: Woodcock-Johnson III normative update*. Rolling Meadows, IL: Riverside.

Miller, J., & Iglesias, A. (2012). Systematic analysis of language transcripts (SALT), version 2012 [Computer software]. Middleton, WI: SALT Software.

National Center for Education Statistics. (2015). *The nation's report card*. Retrieved from http://www.nationsreportcard.gov/reading_math_2015/#reading/acl?grade=4

Nunnally, J. C. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.

Oetting, J. B., & McDonald, J. L. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 207–223. doi:10.1044/1092-4388(2001/018)

Oetting, J. B., & McDonald, J. L. (2002). Methods for characterizing participants' nonmainstream dialect use within studies of child language. *Journal of Speech, Language, and Hearing Research*, *45*, 505–518. doi:10.1044/1092-4388(2002/040)

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Rickford, J. R., & Rickford, R. J. (2000). *Spoken soul: The story of Black English*. New York, NY: Wiley.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, *61*, 1728–1743. doi:10.2307/1130834

Seymour, H. N., Bland-Stewart, L., & Green, L. J. (1998). Difference versus deficit in child African American English. *Language, Speech & Hearing Services in Schools*, *29*(2), 96–108. doi:10.1044/0161-1461.2902.96

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Speece, D. L., Schatschneider, C., Silverman, R., Case, L. P., Cooper, D. H., & Jacobs, D. M. (2011). Identification of reading problems in first grade within a response-to-intervention framework. *Elementary School Journal*, *111*, 585–607. doi:10.1086/659032

Terry, N. P., Connor, C. M., Petscher, Y., & Conlin, C. R. (2012). Dialect variation and reading: Is change in nonmainstream American English use related to reading achievement in first and second grades? *Journal of Speech, Language, and Hearing Research*, *55*, 55–69. doi:10.1044/1092-4388(2011/09-0257)

Terry, N. P., Connor, C. M., Thomas-Tate, S., & Love, M. (2010). Examining relationships among dialect variation, literacy skills, and school context in first grade. *Journal of Speech, Language, and Hearing Research*, *53*, 126–145. doi:10.1044/1092-4388(2009/08-0058)

Terry, N. P., & Scarborough, H. S. (2011). The phonological hypothesis as a valuable framework for studying the relation of dialect variation to early reading skills. In S. Brady, D. Braze, & C. Fowler (Eds.), *Explaining individual differences in reading: Theory and evidence* (pp. 97–117). New York, NY: Taylor & Francis.

Thomas-Tate, S., & Connor, C. M. (2013). *Intervening to support reading comprehension with diverse learners*. Paper presented at the Society for Research on Educational Effectiveness, Washington, DC.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of Word Reading Efficiency–Second Edition*. Austin, TX: PRO-ED.

Troutman, D. E., & Falk, J. S. (1982). Speaking Black English and reading—Is there a problem of interference? *Journal of Negro Education*, *51*(2), 123–133. doi:10.2307/2294677

U.S. Department of Education. (2014). *Thirty-sixth annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2014*. Retrieved from http://www.ed.gov/about/reports/annual/osep

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *Test of Silent Reading Efficiency and Comprehension*. Austin, TX: PRO-ED.

Washington, J. A. (2011). The dialect features of AAE and their importance in LSA. In J. F. Miller, K. Andriacchi, & A. Nockerts (Eds.), *Assessing language production using SALT software: A clinician's guide to language sample analysis* (pp. 113–124). Middleton, WI: SALT Software.

Washington, J. A., Craig, H. K., & Kushmaul, A. J. (1998). Variable use of African American English across two language sampling contexts. *Journal of Speech, Language, and Hearing Research*, *41*, 1115–1124. doi:10.1044/jslhr.4105.1115

Wechsler, D. (2009). *Wechsler Individual Achievement Test, Third Edition*. San Antonio, TX: Pearson.

Wolfram, W. (1998). Language ideology and dialect: Understanding the Oakland Ebonics controversy. *Journal of English Linguistics*, *26*, 108–121. doi:10.1177/007542429802600203

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

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Tests of Achievement, Third Edition*. Itasca, IL: Riverside.

Authors' Note

This research study was supported by a Predoctoral Interdisciplinary Research Training Grant from the Institute of Education Sciences (R305B04074) and a Multidisciplinary Learning Disabilities Center Grant from the National Institute of Child Health and Human Development (P50HD052120). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences, National Institutes of Health, or the National Institute of Child Health and Human Development.

Manuscript received December 2016; accepted July 2017.