



Citation for published version:

Terry, JM, Hendrick, R, Evangelou, E & Smith, RL 2015, Dialect Switching and Mathematical Reasoning Tests: Implications for Early Educational Achievement. in S Lanehart (ed.), The Oxford Handbook of African American Language. Oxford Handbooks, Oxford University Press, pp. 637-658.

Publication date:
2015

Document Version
Peer reviewed version

[Link to publication](#)

University of Bath

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



CHAPTER 34

**DIALECT SWITCHING
AND MATHEMATICAL
REASONING TESTS***Implications for Early Educational Achievement*

J. MICHAEL TERRY, RANDALL HENDRICK,
EVANGELOS EVANGELOU, AND RICHARD L. SMITH

34.1 INTRODUCTION

IN the application of linguistic research to education, a chief focus has been the role that language—more specifically, phonological and grammatical differences between African American English (AAE)¹ and Standard Classroom English (SCE)—may (or may not) play in the academic performance of those African American students who are speakers of AAE. The primary goal of this chapter is to contribute to this growing literature by advancing understanding of the mechanisms by which language affects academic achievement. We identify specific structural features of AAE, whose divergences from SCE, we contend, pose problems by creating a significant additional cognitive load for young AAE speakers who are taught and tested in SCE. We trace and quantify the effect of this load on the scores of AAE-speaking second-grade students on the Woodcock-Johnson-R (hereinafter, WJ-R) Test of Applied Problems (Woodcock and Johnson 1989; for more information, see <http://www.fasttrackproject.org/techrept/w/wjr/>), and we argue that the need to bear its weight may play an important part in preventing a significant number of these students from reaching their full educational potential.

That high levels of AAE use and poor academic performance are correlated has been documented and is widely accepted (Craig, Connor, and Washington 2003; Craig and Washington 2006; Charity, Scarborough, and Griffin 2004; Labov and Baker, this volume); structural explanations for the correlation, such as the one we advance here, however, are less widely embraced. Since the earliest work on the impact of dialectal



differences on learning to read carried out in the 1960s and 1970s, researchers, with some notable exceptions (see Charity et al. 2004; Poe, Burchinal, and Roberts 2004; Labov and Baker, this volume), have generally shifted their thinking from seeking primarily structural explanations for the general relationship between dialect use and academic achievement to more social accounts. William Labov remains one among a few whose work continues to uncover both the social and structural mechanisms that account for the correlation. As he related at a 2011 meeting of the National Research Council, however, even his original research on AAE in South Harlem argued that the main way in which the dialect interferes with school success is its social symbolism as a predictor of academic failure and disciplinary problems. In the wake of such studies, structural accounts tend to be dismissed. For example, highly critical of the view that any structural differences between AAE and SCE are significant enough to explain poor academic performance, sociologist John Ogbu, in his influential article “Beyond Language: Ebonics, Proper English, and Identity in a Black-American Speech Community” (1999), draws warranted attention to the effect that different cultural rules governing dialect use, as opposed to different grammatical rules governing language structure, may have on students mastery of SCE and general academic success. His view is that a major part of current racial disparities in achievement results from many African American students seeing speaking SCE as “talking White,” success in academics as “acting White,” and both the former and the latter as being in conflict with their Black identities (Fordham and Ogbu 1986). Other researchers, for example, Tyson, Darity, and Castellino (2005), argue that in the main, Black students, like their White peers, are achievement oriented, and that the stigma of success in school is generalizable beyond any one group. Their work suggests that school structure, rather than home culture, offers a better explanation for any racialized peer pressure against academic achievement that might exist. Further, they argue that recognizing the similarity between the stigma of “acting White” for Black students and that of “acting high and mighty” for low-income Whites is critical in understanding the issues concerning Black students’ academic success (Tyson, Darity, and Castellino 2005).

No matter the specific mechanisms at work, students’ relationships to their language and other issues of identity undoubtedly affect their acquisition of SCE, and, as a result, their test scores on language-related tasks and academic performance in general. Still, the extent to which differences in the structural features of AAE and SCE themselves may help explain why AAE-speaking children tend to fare poorly in school remains an open question. We believe that finding an answer to this question will require the sorting out of the relative roles that language structure and culture (both at home and at school) play in the process, but that structural mismatches cannot be ignored. In this account, we outline a structural hypothesis aimed at helping explain the correlation between high levels of AAE usage and low levels of academic achievement. The hypothesis, as suggested, is based on the needs of AAE-speaking students to maintain, as well as switch between, grammatical structures from two different dialects, thereby adding a cognitive load to the language processing task. While broad in their potential impact on education, the structural differences we identify as problematic are specific enough to lend themselves to the consideration of practical intervention strategies.

34.2 A STRUCTURAL HYPOTHESIS

34.2.1 Beyond Reading

Students who read well are most apt to experience success in other academic areas, and those who do not usually face wide-ranging academic problems. Because of this, research aimed at determining the role of dialectal variation in the inequitable educational outcomes of African American and White students has, to a great extent, focused on the disparity between the reading abilities of AAE-speaking African American children and their White SCE-speaking peers. The chief hypothesis guiding much of the work in this area has been that differences in AAE and SCE phonology result in a basic sound to written-letter decoding problem for many AAE speakers (see Labov and Baker 2010; Labov and Baker, this volume). Other studies have suggested that the purely phonological differences that exist between the two dialects have far less effect on children's learning to read than is often argued (Harber 1977; Hart, Guthrie, and Winfield 1980; Gemake 1981).

There is mounting evidence that differences in AAE and SCE morphology and syntax may have a greater effect on the process of children's learning to read than differences in phonology (Bartel and Axelrod 1973; Steffensen et al. 1982; Craig and Washington 2004; Van Hofwegen and Stob 2011). Still, the mechanisms by which morphosyntactic differences influence children's reading proficiency remain unclear. No reading-specific cause-effect relation has been discovered, and there is no guarantee that one exists.

Complementing the important body of work that documents the influences AAE has on literacy (Bartel and Axelrod 1973; Steffensen et al. 1982, Purcell-Gates 1996; Gutman, Sameroff, and Eccles 2002; Charity, Scarborough, and Griffin 2004; Craig and Washington 2006; Labov and Baker 2010; Labov and Baker, this volume; Mills and Washington, this volume), we seek to broaden the scope of the discussion on dialect and achievement beyond reading to other critical areas of early education. We test the hypothesis that the morphosyntactic organization of AAE, to the extent that it contrasts with SCE, has significant effects on the performance of AAE-speaking African American second-grade students on the WJ-R Applied Problems subtest, a test of mathematical reasoning. Even though students taking this test are provided with written copy, the test questions are read aloud by the test administrator and repeated as often as a student might require in order to lessen the role that reading likely plays in the process, especially for such young readers. The task, then, raises the question of processing grammatical differences rather than phonological decoding per se.

34.2.2 The Search for a Mechanism

We analyze the performance of young students because we believe it offers the earliest and clearest venue for assessing the differing causal explanations for understanding the correlation between AAE and low academic achievement. One line of explanation

views AAE as principally a series of linguistic markers correlated with social conditions extrinsic to the dialect that negatively impact achievement. For example, Craig and Washington's (2006) analysis holds that AAE's correlation with socioeconomic status, race, teacher expertise, and home literacy habits results from its speakers' limited experience with SCE used in educational settings. Students with this limited experience, it is suggested, face a disadvantage in comparison to students who have more extensive experience with SCE. This general picture is made somewhat more complex by the recognition that some students have linguistic skills that allow them to exploit even limited experience with SCE more efficiently than others. Other extrinsic explanations, as discussed earlier, suggest a determining role for broader cultural practices and styles of interaction in the home (Heath 1983; Fordham and Ogbu 1986; Roberts, Burchinal, and Durham 1999).

A second strategy of explanation attributes the effect to the intrinsic linguistic difference between AAE and SCE. On this view, a mismatch between the organization of the grammars of AAE and SCE poses a burden for children either because they need to switch between the two dialects (Green 1995) or because semantic differences systematically lead them astray (Torrey 1983). Here, too, differences in students' linguistic abilities, whether due to differences in familiarity with SCE or other cognitive skills, complicate the picture. Of course, these two broad approaches are not mutually exclusive, and each likely has a role in a full explanation of the effect AAE has on academic achievement. We believe that understanding in this area will most likely be advanced by greater attention to the relative weight of various explanatory factors, and by greater specificity about the causal mechanisms that could yield observed results.

Factors extrinsic to language can be expected to exert broad influence within a domain of achievement. In contrast, intrinsic factors concerning mismatches between AAE and SCE should have effects that closely track the distribution of those mismatches. Therefore, we test our hypothesis that the morphosyntactic organization of AAE has an effect on AAE-speaking African American students' performance on the WJ-R Applied Problems subtest by determining whether performance on that test correlates with specific structural mismatches between AAE and SCE. We hypothesize that morphemic divergences, in particular, will be the most likely locus for observing the effects of such mismatches. In the end, we find support for an explanation based on the need for children to maintain separate systems and switch between different morphological representations rather than differences in the content carried by those representations.

34.3 THE MATHEMATICAL REASONING DATA

Eighty-seven African American students were recruited from North Carolina community-based childcare centers to participate in a longitudinal study of children's health and development (Roberts et al. 1995). As a part of this study, at regular intervals, language samples were taken from the students, and they were administered a

series of diagnostic tests to assess their linguistic and other cognitive abilities. Of central importance here, these tests included both the Calculation and Applied Problems subtests from the WJ-R Psycho-Educational Battery. Applied Problems is a subtest that assesses skill in analyzing and solving verbal math problems, or “word problems,” as distinct from the Calculation subtest, which tests accuracy of calculation procedures. For example, the question, “If you have eight pennies and you spend two of them, how many pennies would you have left?” has the form of a typical Applied Problems question. In addition to the Applied Problems and Calculation subtests of the WJ-R, other diagnostic tests included the WJ-R Letter-Word Identification subtest, which assesses skills in identifying isolated letters in words, and the Clinical Evaluation of Language Fundamentals (CELF-3) (Semel, Wiig, and Secord 1995), an instrument designed to measure overall receptive and expressive language ability. In addition to these tests, conversational language samples were collected from each of the study participants. Each sample was transcribed using Systematic Analysis of Language Transcripts (SALT) (Miller and Chapman 2000). Study participants were also screened to identify any hearing loss. All tests were administered by one of seven trained examiners with expertise in speech and language assessment. All tests were given at the Frank Porter Graham Child Development Center, a university research facility at the University of North Carolina.

The WJ-R subtests were given at kindergarten entry and at the end of each grade year beginning with first grade. The data analyzed here include the individual responses of seventy-five of the original study participants at second grade. The mean age of the students at the time of testing was 8.32 years. Standard scores, called *W* scores, were calculated from the students’ results. *W* scores are based on the Rasch ability scale (see <http://www.rasch-analysis.com/rasch-analysis.htm> for more explanation) and are centered on a value of 500, which is the approximate average performance of a beginning fifth grader. The sixty questions on the Applied Problems subtest increase in difficulty and are divided into pages. Students were “ceiling tested” by complete pages until the six highest-numbered items were failed, or until the last test item was answered.

In addition to data concerning individual students’ performance on each test question, three members of our team coded each test question for a range of linguistic properties. Reliability between coders was established over the last ten questions, as linguistically speaking, the last questions are the most complex, and therefore most likely to reveal any coding inconsistencies. Subsequently six morphological features were chosen for further statistical analysis. The first five—past tense *-ed*, past participle *-en*, past tense copula *was/were*, auxiliary *have*, and third person singular *-s*—were all chosen because they have been identified as points of divergence between AAE and SCE (Green 2002; Craig and Washington 2006; Wolfram and Shilling-Estes 2006). They may, however, represent different types of divergences (Green 2011). The final feature, counterfactual conditional *if + -ed*, was selected because of its importance to reasoning tasks and the possibility of interaction with *-ed*. All six features are listed along with SCE and AAE examples in table 34.1.

We seek, then, to determine whether the African American students in this study’s performance on the WJ-R Applied Problems subtest correlates with those mismatches

Table 34.1 The Six Morphosyntactic Features Analyzed

Morpheme or Morpheme Combination Coded	Examples	
	Standard Classroom English (SCE)	African American English (AAE)
Past tense <i>-ed</i>	Jill walked Jack to school.	Jill walked Jack to school. OR Jill walk Jack to school.
Participle <i>-en</i>	Jill has written a letter.	Jill have wrote a letter.
Past tense copula ("was," "were")	Jack and Jill were late for school.	Jack and Jill was late for school.
Auxiliary "have"	Jill has written a letter.	Jill have wrote a letter.
Third person singular <i>-s</i>	Jill eats a lot of ice-cream.	Jill eat a lot of ice-cream.
Counterfactual conditional (<i>if + -ed</i>)	If Jill walked Jack to school, he would get there earlier.	If Jill walk Jack to school, he would get there earlier. OR If Jill walk Jack to school, he would get there earlier.

between AAE and SCE that are represented by the linguistic features in table 34.1. Simply being African American, however, does not guarantee that one is a speaker of AAE. Further, there is substantial variation among AAE speakers in the use of those AAE features that contrast with SCE. To measure the variation in students' use of AAE, we calculated dialect density measures (DDMs) from unscripted language samples for the students, using the list of AAE features given in Craig and Washington (2004). A DDM is a measure of the rate of dialect feature production calculated as a ratio of number of dialect features to number of words or utterances (Oetting and McDonald 2002; Craig and Washington 2004; Renn and Terry 2009). The mean DDM score for the students in the study is 0.168 with a standard deviation of 0.124. As our interest here is in quantifying the amount of AAE a student uses rather than labeling him or her as an AAE speaker or SCE speaker, we do not employ a "cut off" score for AAE speaker status. In interpreting these numbers, it is important to recognize that the vast majority of AAE speech overlaps with SCE and other dialects of English.²

34.3.1 Establishing a Linguistic Effect

As previously outlined, our initial data set comprised the individual responses of seventy-five students to the portion of the sixty WJ-R Applied Problems subtest questions that each of them answered. These data were further complemented by counts of the number of times that each of the six linguistic features summarized in

table 34.1 appear in each of the WJ-R Applied Problems subtest questions, and DDMs of each student's AAE production. Our goal here is to determine whether the features in table 34.1 influence students' performance on the WJ-R Applied Problems subtest. To this end, we model whether or not a student answers a given problem correctly as a function of that student's general mathematical ability, the level of difficulty of the problem itself, and the presence or non-presence of any of the linguistic features in table 34.1. We employ a Bayesian Markov Chain Monte Carlo Method (MCMC) to estimate the unknown parameters, including the effect of the features on a student.³ Full details of the model including estimates of model's goodness of fit are given in Appendix A.

We hypothesized that morphemic mismatches between AAE and SCE would affect student performance in learning and testing situations. For each linguistic feature we examined, table 34.2 provides the correlation between a student's total score on the WJ-R Applied Problems subtest and the measure of the influence of that feature on a given student provided by the model. If there is no correlation between a student's score and the linguistic feature, we expect the correlation to be near zero. A high positive or negative value indicates that the effect of the feature is strong.

In the case of a positive correlation, a high negative feature effect on a student indicates that student has worse than average scores on questions in which the feature appears; a high positive feature effect, on the other hand, indicates the student has higher than average scores in which the feature appears. The reverse is true in the case of a negative correlation. Table 34.2 also provides the *p*-value for the null hypothesis—that there is no correlation between a student's score and the linguistic feature—as well as the lower and upper bounds for the 95 percent confidence interval for the correlation.

The results in tables 34.2 and 34.3 suggest that the linguistic features we examined do influence students' overall scores, although the effect of some seems to be negligible.

Table 34.2 Test Score and Feature Effect Correlations

Linguistic feature	Correlation with score	Feature effect statistics		
		95% Confidence interval		
		Lower bound	Upper bound	<i>P</i> -value
Past tense <i>-ed</i>	-0.14	-0.09	0.36	0.23
Participle <i>-en</i>	0.12	-0.11	0.34	0.29
Past tense copula ("was," "were")	0.19	-0.04	0.40	0.10
Auxiliary "have"	0.16	-0.07	0.37	0.18
Counterfactual conditional (<i>if + -ed</i>)	0.43	0.23	0.60	<0.01
Present third singular <i>-s</i>	0.56	0.38	0.70	<0.01

Table 34.3 Feature Effects Averaged across Students

Linguistic feature	Feature effect statistics		
	Correlation with score	Standard deviation	Coefficient of variation
Past tense <i>-ed</i>	-0.03	0.05	-1.66
Participle <i>-en</i>	-0.07	0.29	-3.92
Past tense copula ("was," "were")	-0.08	0.03	-0.45
Auxiliary "have"	-0.10	0.03	-0.03
Counterfactual conditional (<i>if + -ed</i>)	0.12	0.05	0.41
Present 3rd singular <i>-s</i>	0.50	0.68	-1.36

Third person singular *-s* appears to have the greatest effect; in contrast, the past tense copula *was/were* and counterfactual conditional *if + -ed* appear to have the least effect. Although the counterfactual conditional *if + -ed* is estimated to have a small effect, it is noteworthy that the effect it has is facilitative; its presence increases the likelihood that a student will answer a question correctly. As this feature was included in our coding because of the possibility of interaction with *-ed*, a feature whose effect was not facilitative, and not because counterfactuals appear to work differently in AAE than in SCE, this positive effect is likely due to the counterfactual's ability to make transparent the logic of questions in which it is found.

Turning to table 34.3, for each linguistic feature, this table shows the value of the model's measure of the effect of that feature on a student averaged across students, its standard deviation, and coefficient of variation.

While table 34.2 shows that, according to the model, third person singular *-s* has the greatest effect on students' overall scores, in table 34.2 we are able to see that it also exhibits a high degree of variation. This indicates that there are some students who are highly affected by the presence of this particular feature and others who are not. This finding is consistent with those of Johnson, de Villiers, and Seymour (2005) and de Villiers and Johnson (2007).

The data in table 34.2 can also be represented as histograms of the effect of each linguistic feature on the model's measure of the influence of that feature on a given student.⁴ Such histograms are given in figure 34.1. The histograms show clearly the influence of each linguistic feature on the students' scores.

These histograms visually demonstrate that third person singular *-s* has the widest variation with multiple major groups of students, some highly affected by the feature, others moderately, and still others little at all. Past tense copula *was/were* and participle

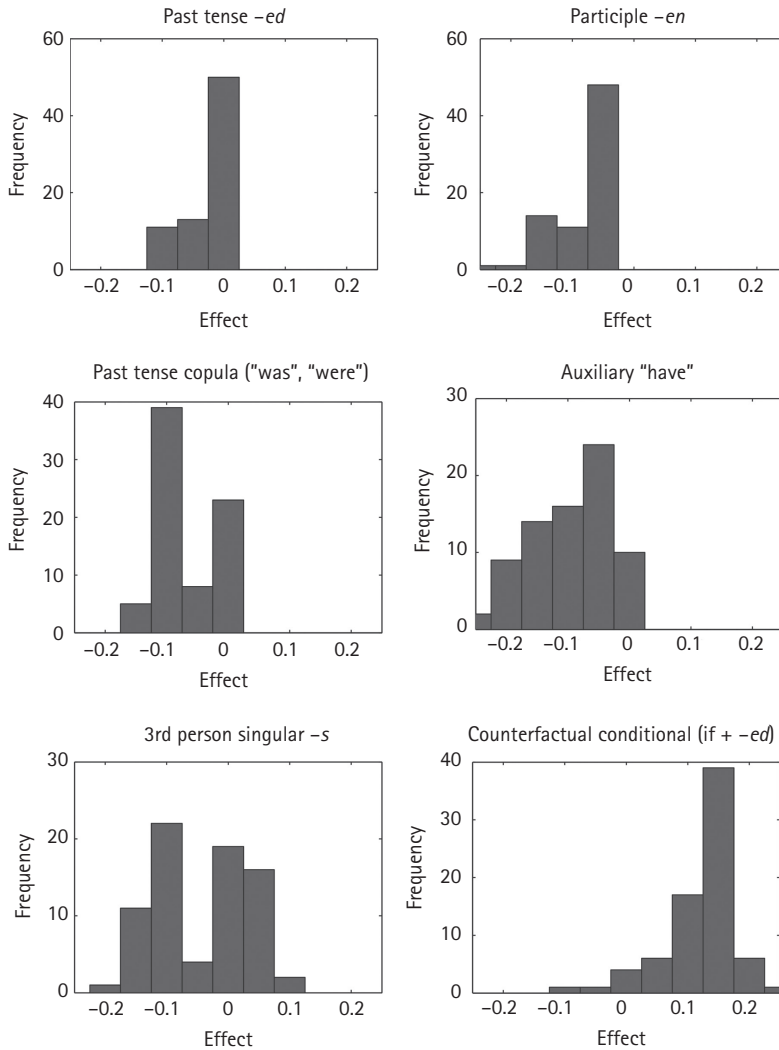


FIGURE 34.1 Estimated effect of linguistic features on student scores.

-en also show at least two groups, although the variation is smaller than that shown by third person singular *-s*. Past tense *-ed* exhibits the least variation.

34.3.2 The Size of the Effect

Our model allows us to estimate the size of a feature's effect on the average student in each of the groups identified. For third person singular *-s*, the most important feature lowering the overall score, the results are as follows. In the highly affected group, roughly

15 percent of the students, the average student would answer 9 percent more questions correctly if the effect of this feature were removed. The size of this effect, then, appears to be educationally significant.

34.3.3 Variation in the Effect on Students

Of the morphemes considered in our study, third person singular *-s* not only had the strongest effect on students, but it also showed the widest variation, splitting the students into three distinct groups: those highly affected by the presence of the feature, those moderately affected, and those who showed little effect. To aid in both theoretical and practical concerns, we would like to know why any morphological feature would affect different AAE speakers as differently as third person singular *-s*.

Two possible reasons unrelated to dialect are available for consideration. One could argue that although this and the other effects we have identified track specific morphological features, the grouping is, in fact, driven by the students' general intelligence rather than the organization or use of their individual grammars. The case for general intelligence has been made based largely on the fact that there is a correlation between scores on tests of demonstrably distinct domains of cognition. For example, while a great deal of evidence shows that verbal memory (and ability) is distinct from spatial memory (and ability), research shows a correlation between subjects' scores on tests of the two (Gardner, Kornhaber, and Wake 1996). If general intelligence or access to general working memory is the hidden driver behind the linguistic effects and groupings that we have identified, we would expect that introducing a measure of spatial reasoning ability—an ability distinct from linguistic abilities but presumably driven by general intelligence—into our MCMC algorithm should lessen the apparent effect of the linguistic features. Using the Block Design subtest from the third edition of the Wechsler Intelligence Scale for Children (Wechsler 1991), as such a measure, we re-ran our MCMC algorithm. This subtest involves copying small geometric designs with four or nine larger plastic cubes. Our re-run of the algorithm showed no significant difference in the size of the effects identified or the groupings of students affected. The independence of spatial reasoning abilities from the linguistic effects is strong evidence that they and the groupings are driven by more linguistically specific factors than general intelligence.

One possible language-based explanation unrelated to dialect might be that the students most affected by the linguistic features are either linguistically delayed or disordered. As the students in this study are beyond the age when problems with morphosyntax typically suggest delay, disorder is the more plausible of these two options. However, it too is unlikely as the student's scores on the Clinical Evaluation of Language Fundamentals (CELF-3) (Semel, Wiig, and Secord 1995), a test commonly used by speech-language pathologists to identify language disorder, do not indicate disorder within the group of students most affected by the linguistic features. Thus, it seems as though a linguistic cause other than delay or disorder is at the root of the variation in

the effect of the linguistic features on students' test performance. Before proposing a plausible cause for this variation, we consider the variation in the effects of the linguistic features themselves.

34.3.4 Variation in the Effects of Features

In explaining why any of the linguistic features we tested would have an effect on test performance, and why some, like third person singular *-s*, have a greater effect than others, our leading hypothesis is that morphosyntactic features whose semantic content is phonologically null in one dialect but not the other pose the greatest difficulty. In this case, SCE third singular *-s* carries present tense meaning while that found in AAE sentences like "John eat" is not phonologically expressed by them.⁵ (See Terry et al. [2010] for further discussion of this perspective and Green [2011] for arguments that third person singular *-s* is not a part of child AAE grammar.)

An additional strength of the general hypothesis is that it may also allow us to explain the variation in the effect this feature has on student performance by drawing a connection to an otherwise anomalous finding noted in Craig and Washington (2004). They report that dialect switching between AAE and SCE is typically accompanied by reduced sentence complexity on the part of the speaker. This suggests to us that dialect switching is purchased at the cost of linguistic complexity. Viewed in this light, the results of our MCMC analysis might suggest that children who must switch between AAE and SCE during mathematical testing sacrifice cognitive resources that would otherwise be available for actual problem solving had they not needed to switch. Although we hypothesize that some morphemes are cognitively taxing in dialect switching, we distinguish ours from the view that the chief source of AAE speakers' problems with third person singular *-s* is a confusion with the meaning of the homophonous plural morpheme (Torrey 1969). Likewise, we distinguish our dialect-shifting hypothesis from the highly contested view that the semantic organization of AAE is such that it does not allow for the efficient representation of key mathematical concepts (Orr 1987; cf. O'Neil 1990 and Baugh 1999).

34.3.5 Dialect Switching

Without a direct measure of the dialect switching abilities of the students in this study, we are unable to completely confirm our suspicion that it is precisely those students who find it most difficult to switch from AAE to SCE who, in turn, are the most affected by dialect mismatches such as the presence of third person singular *-s* in test questions. We expect, however, that there is considerable overlap between those students who struggle to switch dialects and those students who have the greatest need to do so. Put another way, although all AAE speakers dialect shift to some degree, we expect that the more monodialectal AAE speakers in our sample will, for the most part, be the "heavier"

dialect speakers and vice versa. Given this reasoning, we predict that introducing the students' DDMs into our MCMC algorithm will change the average student value of our measure to the extent to which a given linguistic feature affects a student, its correlation with test scores, and the coefficient of variation for each of the features we considered. The values for these statistics after re-running our algorithm with the DDMs included are given in tables 34.4 and 34.5.

Table 34.4 Feature Effects Averaged across Students (AAE Production Included in Model)

Linguistic feature	Feature effect statistics		
	Mean	Standard deviation	Coefficient of variation
Past tense <i>-ed</i>	-0.03	0.05	-1.40
Participle <i>-en</i>	-0.14	0.32	-0.23
Past tense copula ("was," "were")	-0.20	0.04	-0.19
Auxiliary "have"	-0.16	0.37	-0.23
Counterfactual conditional (<i>if + -ed</i>)	0.11	0.05	0.44
Present third singular <i>-s</i>	0.06	0.07	-1.16

Table 34.5 Test Score and Feature Effect Correlations (AAE Production Included in Model)

Linguistic feature	Feature effect statistics			
	Correlation with score	95% Confidence interval		<i>P</i> -value
		Lower bound	Upper bound	
Past tense <i>-ed</i>	0.14	-0.09	0.36	0.24
Participle <i>-en</i>	0.11	-0.12	0.33	0.35
Past tense copula ("was," "were")	0.18	-0.05	0.39	0.13
Auxiliary "have"	0.14	-0.09	0.36	0.23
Counterfactual conditional (<i>if + -ed</i>)	0.42	0.21	0.59	<0.01
Present third singular <i>-s</i>	0.72	0.59	0.81	<0.01

Across students there is variation in the effect of each of the linguistic features we examined, variation in how well they performed on the test, and variation in their AAE production. With the addition of a DDM term into our model, the correlation between the model's measure of the influence of a given feature on a given student and student test scores, as shown in table 34.4, can be viewed as a measure of the extent to which the variation on those three dimensions overlaps. The strong correlation and high coefficient of variation reported in table 34.4 suggest that the linguistic effects that we have identified are, in fact, effects of dialect and that the stronger the dialect, the stronger the effect. Thus, these numbers are supportive of our dialect-switching hypothesis.

On the other hand, a potential criticism of the dialect-switching hypothesis is that a similar pattern of results might be achieved simply by looking for a correlation between the length of a question (in terms of the number of sentences it contains), its difficulty and students' performance. From this perspective, the number of times third person singular *-s* occurs in a test question might be thought of as a stand in for the number of sentences in that question. While this alternative seems plausible enough on the surface, it does not fit well with the data in our study. This is because it is only narrowly consistent with the information about third singular *-s*; other features, such as the past tense marker *-ed*, or the counterfactual conditional *if + -ed* pattern with multiple sentences in the WJ-R Applied Problems subtest, yet they do not show nearly as strong correlation with students' performance. And in the case of the counterfactual conditional *if + -ed*, the correlation runs in the opposite direction: the more instances of third person singular *-s* there are in a question the worse students do; the more instances of the counterfactual conditional there are, the better. Patterns such as these show the value of an analysis that is more finely grained than simply counting sentences, one that focuses instead on the particular morphemes that mismatch in AAE and SCE within multiple sentences.

Our dialect-switching hypothesis explains the apparent negative effect of AAE on WJ-R Applied Problems subtest that we have identified in cognitive processing terms as it places the source of the effect in AAE-speaking students' need to switch between two dialects with dissimilar morphosyntactic systems. While this is our leading hypothesis, a non-processing reading of the data is, however, still possible. One might argue that lack of familiarity with the narrative style used in the test questions is the source of the difficulty for AAE-speaking students. In order to explore the viability of this line of causal explanation, we investigated the distribution of third person singular *-s* and conditional counterfactual *if + -ed* clauses in the WJ-R test questions. Neither feature appeared to occupy an especially salient position that might affect test performance directly or connect to any known differences in AAE and SCE narrative styles (Champion 2003; Champion and McCabe, this volume). While we cannot rule out the possibility of a successful narrative-based account of the data, before such an explanation can be tested, the narrative features thought to be responsible of the patterns in the data would need to be characterized and the way in which they could interact with question difficulty would need to be made precise. Until such an account is proposed, in our view, the dialect-switching hypothesis remains the clearest explanation with the greatest empirical support.

34.4 CONCLUSION

The results of our analysis show that linguistic features in general, and linguistic features associated with structural differences between AAE and SCE, in particular, can have a significant impact on young AAE-speaking students' performance on tests of mathematical reasoning that are given in SCE. That impact can be facilitative, as in the case of the counterfactual conditional *if + -ed*, a feature associated with reasoning tasks, but not dialectal difference, or it can be inhibitory, as in the case of third person singular *-s*, a point of divergence between the two dialects. Importantly, the inhibitory impact we identify is independent of individuals' abilities in spatial reasoning and, therefore, appears not to be a matter of general intelligence. Nor does it appear to be associated with any language delay or disorder. Instead, the impact that we have documented provides support for our initial hypothesis that some dialectal differences pose problems for some AAE-speaking students on verbally mediated reasoning tasks such as the WJ-R Applied Problems subtest due to the demands of switching between different linguistic representations.

The significance of this line of inquiry is quite broad, having strong implications for both linguistic theory and educational practice. With respect to linguistic theory, we see an important hypothesis that follows from trying to understand why dialect switching poses a cognitive load. Our finding that third person singular *-s* has a significant impact on performance on mathematical reasoning tasks makes the most sense if we tie this impact to differences in whether a linguistic feature has an overt morphemic representation in one or both dialects that must be managed.

With respect to educational practice, the features we identify as inhibiting student performance—third person singular *-s* chief among them—are specific enough to lend themselves to very focused intervention strategies. The ability to arm teachers with the knowledge of which dialectal differences are likely to pose significant problems for learners and which are not holds with it the promise of more targeted strategies for helping AAE-speaking students to navigate the dialects that are used at home and at school, and for reducing any negative effects that differences between them might cause. Targeting, however, is only half of the issue. Understanding what makes a feature like third person singular *-s* more problematic than other features is important in determining what type of targeted intervention will be most effective. Thus, the practical issue of intervention is very much connected to the more theoretical issues outlined above.

There is no simple answer to the question of how dialectal difference affects educational achievement. No doubt a variety of complex social and structural factors have roles they play and the results of our analysis argue that differences in the morphosyntactic inventories of dialects and the need to manage them deserve attention as one of those factors.

APPENDIX A

A.1 The Model

Let z_{ij} denote the score of student i on question j ; z_{ij} takes the value 1 for a correct answer and the value 0 for an incorrect answer. Students in this study were ceiling tested, and no student was asked every question. We treat unasked questions as missing values and ignore them. Define y_{ij} as a measure of how well the student i knows the answer to question j and treat it as an unobserved random variable such that y_{ij} is an unobserved random variable such that $z_{ij} = 1$ if $y_{ij} > 0$ and $z_{ij} = 0$ if $y_{ij} \leq 0$. Our principal interest lies in the effect that each feature k has on a student i , represented as α_{ik} . We let x_{jk} represent the number of times the linguistic feature k appears in a question j and use it as a measure of the influence of k on the question. In addition to the six features we study, we expect a student's answer to be affected by that student's overall mathematical ability and the difficulty of the question being asked. We represent these effects as η_i and β_j , respectively.

The model we use is:

$$y_{ij} = \eta_i + \beta_j + \sum_{k=1}^6 \alpha_{ik} x_{jk} + \epsilon_{ij} \quad (1)$$

for $i = 1, \dots, 75$ and $j = 1, \dots, 60$.

The term ϵ_{ij} represents the error of the model. This error is assumed to be introduced by other factors we have not taken into account (e.g., socioeconomic status and properties of the student's home environment). It is taken to be independent of the other variables.

A.2 Assumptions

Because the students and the questions were chosen randomly from a larger group of students and questions, it is logical to treat all η_i , β_j , and α_{ik} as random effects.

We assume that random effects have a normal distribution and that they are independent of each other:

$$\eta_i \sim N(\mu_\eta, \sigma_\eta^2), \quad i = 1, \dots, 75$$

$$\beta_j \sim N(\mu_\beta, \sigma_\beta^2), \quad j = 1, \dots, 60,$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2), \quad i = 1, \dots, 75, \quad j = 1, \dots, 60,$$

$$\alpha_{ik} \sim N(v_k, \tau_k^2), \quad i = 1, \dots, 75, \quad k = 1, \dots, 6.$$

The parameters μ_η , μ_β , v_k , σ_η^2 , σ_β^2 , σ_ϵ^2 , and τ_k^2 are unknown and, together with η_i , β_j , and α_{ik} , must be estimated from the data.

A.3 Approach

We apply a Bayesian Markov Chain Monte Carlo (MCMC) method (Young and Smith 2005, 22–48) to estimate the unknown parameters.

A.3.1 Simplification of the Model

Given the initial assumptions we can simplify the model:

- The model is only affected by the difference between μ_η and μ_β and not by their individual values. So we set $\mu_\beta = 0$.
- If we multiply all the y_{ij} 's by the same positive constant, the values of z_{ij} do not change, so we can set $\sigma_\epsilon^2 = 1$.

The model becomes:

$$\begin{aligned} \eta_i &\sim N(\mu, \kappa_\eta^{-1}), \\ \beta_j &\sim N(0, \kappa_\beta^{-1}), \\ \epsilon_{ij} &\sim N(0, 1), \\ \alpha_{ik} &\sim N(v_k, \lambda_k^{-1}), \\ y_{ij} &= \eta_i + \beta_j + \sum_{k=1}^6 \alpha_{ik} x_{jk} + \epsilon_{ij}, \\ z_{ij} &= \begin{cases} 1 & \text{if } y_{ij} > 0, \\ 0 & \text{if } y_{ij} \leq 0. \end{cases} \end{aligned}$$

Note that we now write the variances as κ_η^{-1} , κ_β^{-1} , and λ_k^{-1} instead of σ_η^2 , σ_β^2 , and τ_k^2 to simplify calculations.⁶

A.3.2 Priors for the Hyperparameters

The parameters μ , v_k , κ_η , κ_β , and λ_k are called hyperparameters.

We used:

$$\begin{aligned} \mu &\sim U(-\infty, \infty), \\ v_k &\sim U(-\infty, \infty), \\ \kappa_\eta &\sim \text{Gamma}(a, b), \\ \kappa_\beta &\sim \text{Gamma}(a, b), \\ \lambda_k &\sim \text{Gamma}(a, b), \end{aligned}$$

where U denotes the uniform distribution and Gamma ⁷ the gamma distribution and $a = b = 0.01$.⁸

We assume that the hyperparameters are independent from each other and also they are independent from the effects α , β , and η and from the error ϵ .

A.4 The MCMC Algorithm

Let I represent the number of students ($I = 75$). J is the number of questions ($J = 60$), and K is the number of linguistic factors ($K = 6$). The joint density of $(\kappa_\eta, \lambda_k, \mu, \nu_k, \eta_i, \beta_j, \alpha_{ik}, y_{ij}, z_{ij})$ is proportional to

$$\begin{aligned} & \kappa_\eta^{a-1} e^{-b\kappa_\eta} \cdot \kappa_\beta^{a-1} e^{-b\kappa_\beta} \cdot \prod_{k=1}^K \left\{ \lambda_k^{a-1} e^{-b\lambda_k} \right\} \cdot \prod_{i=1}^I \kappa_\eta^{\frac{1}{2}} e^{-\frac{1}{2}\kappa_\eta(\eta_i - \mu)^2} \cdot \prod_{j=1}^J \kappa_\beta^{\frac{1}{2}} e^{-\frac{1}{2}\kappa_\beta\beta_j^2} \\ & \prod_{i=1}^I \prod_{k=1}^K \lambda_k^{\frac{1}{2}} e^{-\frac{1}{2}\lambda_k(\alpha_{ik} - \nu_k)^2} \cdot \prod_{i=1}^I \prod_{j=1}^J e^{-\frac{1}{2}(y_{ij} - \eta_i - \beta_j - \sum_{k=1}^K \alpha_{ik}x_{jk})^2} \cdot Q(y_{ij}, z_{ij}) \end{aligned} \quad (2)$$

where

$$Q(y, z) = \begin{cases} 1 & \text{if } y > 0 \text{ and } z = 1, \\ 1 & \text{if } y \leq 0 \text{ and } z = 0 \\ 0 & \text{otherwise.} \end{cases}$$

With the exception of z_{ij} , all the variables in (2) are unknown. The Bayesian solution to this problem is to construct the conditional density of $(\kappa_\eta, \kappa_\beta, \lambda_k, \mu, \nu_k, \eta_i, \beta_j, \alpha_{ik}, y_{ij})$ given all the z_{ij} . The basic idea of MCMC sampling is to construct a Monte Carlo sample from the joint density (2) by successively updating each of the unknown random variables.

A.5 Updating Scale Parameters

The scale parameters are κ_η , κ_β , and λ_k , $k = 1, \dots, K$. Updating the scale parameters consists of a random sample of one observation from the $\text{Gamma}(a', b')$ distribution where:

- for κ_η , $a' = a + \frac{1}{2}I$ and $b' = b + \frac{1}{2}\sum_i (\eta_i - \mu)^2$
- for κ_β , $a' = a + \frac{1}{2}J$ and $b' = b + \frac{1}{2}\sum_j \beta_j^2$
- for λ_k , $a' = a + \frac{1}{2}I$ and $b' = b + \frac{1}{2}\sum_i (\alpha_{ik} - \nu_k)^2$

A.6 Updating Location Parameters

The location parameters are μ , ν_k , η_i , β_j , and α_{ik} . Updating the location parameters consists of a random sample of one observation from the $N\left(\frac{B}{A}, \frac{1}{A}\right)$, where:

- for μ , $A = I\kappa_\eta$, $B = \kappa_\eta \sum_i \eta_i$
- for ν_k , $A = I\lambda_k$, $B = \lambda_k \sum_i \alpha_{ik}$



- for η_i , $A = \kappa_\eta + J$, $B = \mu\kappa_\eta + \sum_j (y_{ij} - \beta_j - \sum_k \alpha_{ik} x_{jk})$
- for β_j , $A = \kappa_\beta + I$, $B = \sum_i (y_{ij} - \eta_i - \sum_k \alpha_{ik} x_{jk})$
- for α_{ik} , $A = \lambda_k \sum_j x_{jk}^2$, $B = \lambda_k \nu_k \sum_j x_{jk} (y_{ij} - \eta_i - \sum_{k'} \alpha_{ik'} x_{jk'})$

A.7 Updating y_{ij}

The conditional distribution of y_{ij} given all the other unknowns is $N(\eta_i + \beta_j + \sum_k \alpha_{ik} x_{jk}, 1)$ (including the condition $Q(y_{ij}, z_{ij}) = 1$). Rejection sampling to sample y : consecutive values were generated from the conditional distribution until the condition $Q(y_{ij}, z_{ij}) = 1$ is satisfied.

A.8 Implementation

For starting values, we set $y_{ij} = 1$ when $z_{ij} = 1$ and $y_{ij} = -1$ when $z_{ij} = 0$. We set all the location parameters equal to 0 and all the scale parameters equal to 1. We then ran 10,000 iterations as “burn in” updating all the unknowns. The results were discarded. This is done so that the starting values that we chose for the first step would not affect the results. 100,000 more iterations were then carried out, and the results of each 100th step were preserved to compile a sample size of 1,000 from the posterior distributions of the unknown variables. We use the superscript (n) to refer to the n th observation in the sample so that $\alpha_{3,10}^{(45)}$ means the 45th observation in the sample of the parameter $\alpha_{3,10}$.

A.9 Checking the Fit of the Model

We can use at least two different methods to check how well our model explains the data.

A.9.1 Using the Estimated Values, \hat{z}_{ij} Compared to the Original Values of z_{ij}

Using the simulated data, we calculated the values $\hat{z}_{ij}^{(n)}$, where $\hat{z}_{ij}^{(n)}$ is the estimated value of z_{ij} for the n th observation.

First, we calculate $\hat{y}_{ij}^{(n)}$ by:

$$\hat{y}_{ij}^{(n)} = \eta_i^{(n)} + \beta_j^{(n)} + \sum_{k=1}^6 \alpha_{ik}^{(n)} x_{jk}$$

where $\eta_i^{(n)}$, $\beta_j^{(n)}$ and $\alpha_{ik}^{(n)}$ refer to the n th observation in the sample. Then, we set $\hat{z}_{ij}^{(n)} = 1$ if $\hat{y}_{ij}^{(n)} > 0$ and $\hat{z}_{ij}^{(n)} = 0$ if $\hat{y}_{ij}^{(n)} \leq 0$. This is done for each n . Then for each pair (i, j) we calculate the sample mean \hat{z}_{ij} of $\hat{z}_{ij}^{(n)}$, $\hat{z}_{ij} = \frac{1}{1000} \sum_n \hat{z}_{ij}^{(n)}$. This is a number between 0 and 1. We then divide the interval $[0, 1]$ into $L = 10$ equally spaced subintervals: $[0.0, 0.1)$, $[0.1, 0.2)$, \dots , $[0.9, 1.0]$ and I take the



average $\hat{z}_{ij}^{[l]}$ of the \hat{z}_{ij} 's that belong to the subinterval l , $l = 1, \dots, L$. This defines a set of pairs (i, j) .

We also take the average $\hat{z}_{ij}^{[l]}$, of the observed z_{ij} 's, for those (i, j) 's.

We expect that if we plot the $\hat{z}_{ij}^{[l]}$'s against the $\hat{z}_{ij}^{[l]}$'s, then we will get a straight line. The plot is shown in figure A.1. The correlation is 0.8702.

A.9.2 Plots of the Median of β_j 's against the Proportion of Correct Answers for Question j

The overall difficulty of question j is estimated by β_j . We expect an increasing pattern between the median of β_j 's and the average number of correct answers for each question.

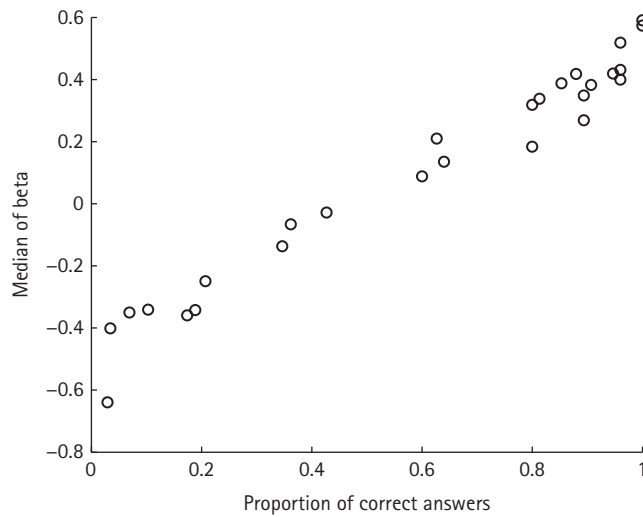


FIGURE A.1 Plot of $\hat{z}_{ij}^{[l]}$'s against $\hat{z}_{ij}^{[l]}$'s.

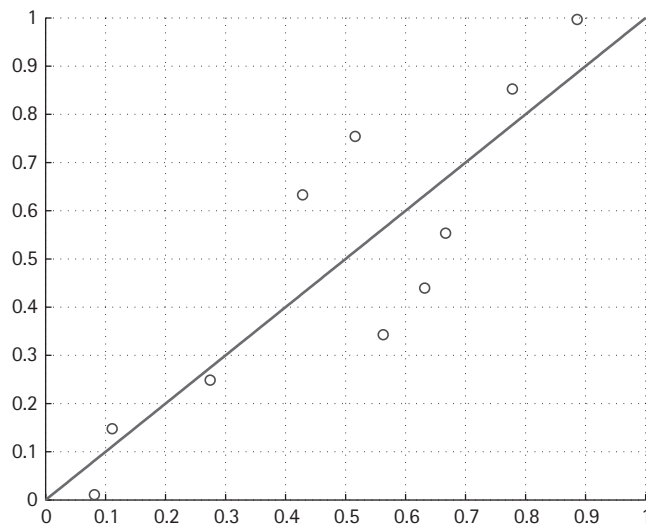


FIGURE A.2 Plot of β against question average correct answers.

The plot is shown in figure A.2. The increasing relationship is more obvious here. The correlation is 0.9818.

NOTES

1. By African American English, we mean the relatively uniform variety spoken by many but not all African Americans throughout the United States. Defined by its grammar and use, we are most concerned with those features of the variety that are common to many of its regional sub-varieties.
2. This overlap highlights a potential problem with the use of DDMs and similar token-based measures of dialect use. Such measures typically count as features of the dialect only those features that contrast with those of more mainstream dialects. See Green (2011) for further discussion of this issue.
3. We treat the model parameters (the student's general mathematical ability, the overall difficulty of the question, and the extent to which the student is affected by a given linguistic feature) as random effects. The Bayesian-MCMC approach to fitting the model is to define prior distributions for these parameters and use Gibbs and Metropolis sampling to construct posterior distributions for all the unknowns. The posterior distributions are then used to determine which, if any, linguistic features influence students' scores. If one treats the influence of the linguistic features as fixed rather than random effects (i.e., if one assumes the influence of a particular linguistic feature is the same for all students), it is possible to estimate the model by standard logistic regression. However, not only is the number of unknown parameters too large to assume fixed effects, but doing so prevents testing whether the amount of AAE a student uses correlates with the effect AAE features have on test performance, an important hypothesis pursued here.
4. One observation of the first linguistic feature, past tense *-ed*, was omitted as an outlier. The value was -0.4 .
5. At first blush, this might appear to apply equally well to past tense *-ed*, a feature that showed little effect, as it does to third person singular *-s*, the feature that showed the greatest effect; past tense *-ed* is often omitted in AAE. However, following Green (2011), we assume such omission is due to the presence of a variable rule within AAE grammar as opposed to the overt marking of the feature not being a part of the grammar as we believe is the case with third person singular *-s*.
6. Because using a *Gamma* distribution for the prior of κ_η , κ_β , and λ_k gives a *Gamma* posterior distribution for those variables so I will get a distribution in a closed form.
7. The density for the *Gamma*(a, b) distribution is $f(x) = \frac{1}{r(a)} b^a x^{a-1} e^{-bx}$.
8. This is a typical choice for the MCMC.

REFERENCES

- Bartel, N. R., and J. Axelrod. 1973. "Nonstandard English Usage and Reading Ability in Black Junior High Students." *Exceptional Children* 38: 653–55.
- Baugh, John. 1999. *Out of the Mouths of Slaves: African American Language and Educational Malpractice*. Austin, TX: University of Texas.



- Champion, Tempii B. 2003. *Understanding Storytelling among African American Children: A Journey from Africa to America*. Mahwah, NJ: Lawrence Erlbaum.
- Charity, Anne H., Hollis S. Scarborough, and Darion M. Griffin. 2004. "Familiarity with 'School English' in African American Children and Its Relation to Early Reading Achievement." *Child Development* 75: 1340–56.
- Craig, Holly K., Carol M. Connor, and Julie A. Washington. 2003. "Early Positive Predictors of Later Reading Comprehension for African American Students: A Preliminary Investigation." *Language, Speech, and Hearing Services in the Schools* 34: 31–43.
- Craig, Holly K., and Julie A. Washington. 2004. "Grade-Related Changes in the Production of African American English." *Journal of Speech, Language and Hearing Research* 47: 450–63.
- . 2006. *Malik Goes to School: Examining the Language Skills of African American Students from Preschool–5th Grade*. Mahwah, NJ: Lawrence Erlbaum.
- de Villiers, Jill G., and Valerie E. Johnson. 2007. "The Information in Third-Person /s/: Acquisition across Dialects of American English." *Journal of Child Language* 34: 113–58.
- Fordham, Signithia, and John Ogbu. 1986. "Black Students' School Success: Coping with the 'Burden of Acting White.'" *Urban Review* 18: 176–206.
- Gardner, Howard, Mindy L. Kornhaber, and Warren K. Wake. 1996. *Intelligence: Multiple Perspectives*. Fort Worth, TX: Harcourt Brace College Publishers.
- Gemake, Josephine S. 1981. "Interference of Certain Dialect Elements with Reading." *Reading Improvement* 18: 183–99.
- Green, Lisa J. 1995. "Study of Verb Classes: Implications for Education." *Linguistics in Education* 7: 65–81.
- . 2002. *African American English: A Linguistic Introduction*. Cambridge: Cambridge University.
- . 2011. *Language and the African American Child*. Cambridge: Cambridge University.
- Gutman, Leslie M., Arnold J. Sameroff, and Jacquelynne S. Eccles. 2002. "The Academic Achievement of African American Students during Early Adolescence: An Examination of Multiple Risk, Promotive, and Protective Factors." *American Journal of Community Psychology* 30: 367–99.
- Harber, Jean R. 1977. "Influence of Presentation Dialect and Orthographic Form on Reading Performance of Black Inner-City Children." *Educational Research Quarterly* 2: 9–16.
- Hart, Jane T., John T. Guthrie, and Linda Winfield. 1980. "Black English Phonology and Learning to Read." *Journal of Educational Psychology* 72: 636–46.
- Heath, Shirley Brice. 1983. *Ways with Words: Language, Life, and Work in Communities and Classrooms*. Cambridge: Cambridge University.
- Johnson, Valerie E., Jill G. de Villiers, and Harry N. Seymour. 2005. "Agreement without Understanding? The Case of Third Person Singular /s/." *First Language* 25: 317–30.
- Labov, William, and Bettina Baker. 2010. "What is a Reading Error?" *Applied Psycholinguistics* 31: 735–57.
- Miller, Jon F., and Chapman, Robin. 2000. "SALT: Systematic Analysis of Language Transcripts (Version 9.0)." Madison: Language Analysis Laboratory, Waisman Center, University of Wisconsin.
- O'Neil, Wayne. 1990. "Dealing with Bad Ideas: Twice Is Less." *English Journal* 79: 80–88.
- Oetting, Janna B., and Janet L. McDonald. 2002. "Methods for Characterizing Participants' Nonmainstream Dialect Use in Child Language Research." *Journal of Speech, Language, and Hearing Research* 45: 505–18.



- Ogbu, John. 1999. "Beyond Language: Ebonics, Proper English, and Identity in a Black-American Speech Community." *American Educational Research Journal* 36: 147–84.
- Orr, Eleanor Wilson. 1987. *Twice as Less: Black English and the Performance of Black Students in Mathematics and Science*. New York: Norton.
- Poe, Michele D., Margaret R. Burchinal, and Joanne E. Roberts. 2004. "Early Language and the Development of Children's Reading Skills." *Journal of School Psychology* 42: 315–32.
- Purcell-Gates, Victoria. 1996. "Stories, Coupons, and the TV Guide: Relationships between Home Literacy Experiences and Emergent Literacy Knowledge." *Reading Research Quarterly* 31: 406–28.
- Renn, Jennifer, and J. Michael Terry. 2009. "Operationalizing Style: Quantifying Style Shift in the Speech of African American Adolescents." *American Speech* 84: 367–90.
- Roberts, Joanne E., Margaret R. Burchinal, and Meghan Durham. 1999. "Parents' Report of Vocabulary and Grammatical Development of African American Preschoolers: Child and Environmental Associations." *Child Development* 70: 92–106.
- Roberts, Joanne E., Margaret R. Burchinal, Lynn P. Medley, Susan A. Zeisel, Martha Mundy, Jackson Roush, Stephen Hooper, Donna Bryant, and Frederick W. Henderson. 1995. "Otitis Media, Hearing Sensitivity, and Maternal Responsiveness in Relation to Language during Infancy." *Journal of Pediatrics* 126: 481–89.
- Semel, Elenor, Elisabeth H. Wiig, and Wayne A. Secord. 1995. *Clinical Evaluation of Language Fundamentals 3 (CELF3)*. San Antonio, TX: The Psychological Corporation.
- Steffensen, Margaret S., Ralph E. Reynolds, Erica McClure, and Larry F. Guthrie. 1982. "Black English Vernacular and Reading Comprehension: A Cloze Study of Third, Sixth, and Ninth Graders." *Journal of Reading Behavior* 14: 285–98.
- Terry, J. Michael, Randall Hendrick, Evangelos Evangelou, and Richard L. Smith. 2010. "Variable Dialect Switching among African American Children: Inferences about Working Memory." *Lingua* 120: 2463–75.
- Torrey, Jane. 1983. "Black Children's Knowledge of Standard English." *American Educational Research Journal* 20: 627–43.
- Tyson, Carolyn, William Darity Jr., and Domini R. Castellino. 2005. "It's Not a Black Thing: Understanding the Burden of Acting White and Other Dilemmas of High Achievement." *American Sociological Review* 70: 582–605.
- Van Hofwegen, Janneke, and Reuben Stob. 2011. "A Longitudinal Analysis of the Relationship between Reading and AAE Vernacularity." Paper presented at New Ways of Analyzing Variation 40, Washington, DC, October 27–30.
- Washington, Julie A., and Holly K. Craig. 2001. "Reading Performance and Dialectal Variation." In *Literacy in African American communities*, edited by Joyce L. Harris, Alan G. Kamhi, and Karen E. Pollock, 147–68. Mahwah, NJ: Lawrence Erlbaum.
- Wechsler, David. 1991. *The Wechsler Intelligence Scale for Children—Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wolfram, Walt, and Natalie Schilling-Estes. 2006. *American English: Dialects and Variation*. Malden, MA: Blackwell.
- Woodcock, Richard W., and Mary Bonner Johnson. 1989. *Woodcock-Johnson Psycho-Educational Battery—Revised*. Allen, TX: DLM Teaching Resources.
- Young, G. A., and Richard L. Smith. 2005. *Essentials of Statistics Inference*, 22–48. Cambridge: Cambridge University.