Child Development, September/October 2004, Volume 75, Number 5, Pages 1340–1356

Familiarity With School English in African American Children and Its Relation to Early Reading Achievement

Anne H. Charity, Hollis S. Scarborough, and Darion M. Griffin

For children whose everyday speech differs greatly from the School English (SE) they encounter in academic materials and settings, it was hypothesized that greater familiarity with SE would be associated with more successful early reading acquisition. Sentence imitation and reading skills of 217 urban African American students in kindergarten through second grade (ages 5 to 8 years) were assessed. Children in each grade varied widely in the extent to which their imitations of SE sentences included phonological and grammatical forms that are acceptable in African American Vernacular English but not in SE. Higher familiarity with SE (reproducing SE features more often when imitating) was associated with better reading achievement, and these relationships were independent of memory ability.

As they learn to map their knowledge of spoken language onto print, all beginning readers will encounter some disparities between the oral and written representations of phonological, morphosyntactic, and lexical forms. Such opacity will be relatively infrequent for native English speakers whose dialects more closely resemble School English (SE), the predominant dialect in written English and often in classroom instruction. When a child's dialect differs more substantially from SE, however, there will be many more mismatches between oral and written forms, potentially impeding the child's progress in learning to read. One such dialect is African American Vernacular English (AAVE). We therefore hypothesized that greater familiarity with SE would be associated with more successful reading acquisition by young African American students from homes and communities in which AAVE is often spoken.

The phonological and grammatical features that distinguish AAVE (which has also been termed *Black English* and *Ebonics*) from more standard American dialects have been well documented in sociolinguistic studies over the past half century, and detailed descriptions of AAVE features can be found in many sources (e.g., Labov, 1972; Mufwene, Rickford, Bailey, & Baugh, 1998; Rickford, 1999; Washington & Craig, 1994; Wolfram & Fasold, 1974). For example, omission of final consonants (e.g., /bæ/ in AAVE vs. /bæd/ in SE for the word *bad*), reduction of final consonant clusters (e.g., /mos/ vs. /most/ for *most*, /hɛp/ vs. /hɛlp/ for *help*), and the use of stops or labiodental fricatives where SE has interdental fricatives (e.g., /dɪs/ vs. /ðɪs/ for *this*, /bof/ or /bot/ vs. /boθ/ for *both*) are phonologically acceptable in AAVE but not in SE. Absence of grammatical inflections (including plural, possessive, and past tense affixes) and of auxiliary/copular verbs is also sometimes permissible in AAVE, as are various alternative constructions involving negation, pronoun case, and other morphosyntactic elements. Note that what distinguishes AAVE from SE is the frequency with which certain characteristic features occur rather than their sheer presence or absence.

It is now generally acknowledged that, in contrast to earlier views (e.g., Bereiter, 1966), AAVE is not an impoverished version of more standard English dialects. Rather, it is a separate but equivalent system, as complex and rule governed as SE but with some alternative rules and conventions for expressing the

Anne H. Charity, Linguistics Department, University of Pennsylvania and Haskins Laboratories; Hollis S. Scarborough, Haskins Laboratories; Darion M. Griffin, American Federation of Teachers.

This research was sponsored in part by Program Project Grant HD01994 from the National Institute of Child Health and Human Development to Haskins Laboratories and by Grant R215U990010 from the U.S. Department of Education Office of Educational Research and Improvement to the American Federation of Teachers (AFT). We thank Marsha Berger for her cooperation and leadership during her tenure as director of the AFT project. We are also very grateful to Sharon Hughes, Adrienne Dowden, Doris Taylor, Susan O'Donnell, Marie Hotard, Leslie Garibaldi, Joan Dorsey, Lisa Ratstatter, and Benita Nicholson for collecting the data, and to Brook Swainson for assistance with scoring the sentence imitation measure. Valuable input on the research was provided by our colleagues William Labov, Anne Fowler, and Alice Faber. A preliminary report of this study was presented to the Society for the Scientific Study of Reading.

Correspondence concerning this article should be addressed to Anne H. Charity or Hollis S. Scarborough, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695. Electronic mail may be sent to acharity@babel.ling.upenn.edu or hscarborough@att.net.

same syntactic relationships and semantic content (Labov, 1972, 1998; Turner, 1949). More frequent production of AAVE features has been associated with a variety of social factors, including the individual's degree of exposure to speakers of other dialects (Labov & Harris, 1986; Sims, 1982) and the gender and socioeconomic status (SES) of the speaker (Dillard, 1972; Washington & Craig, 1998; Wolfram & Fasold, 1974). There is also variation within individuals in the frequency with which AAVE features are produced across a variety of social settings, and this variation is also related to SES (Baugh, 1983; Washington, Craig, & Kushmaul, 1998).

The notion that reading achievement might be associated with dialect differences is not a new one. Indeed, much of the early research on AAVE was motivated by the concern that the differences between AAVE and SE might be responsible, at least in part, for the reading difficulties of many urban African American children (Labov, 1995; Rickford & Rickford, 1995). In addressing this question, however, scholars have disagreed about the mechanism(s) by which a student's use of AAVE might contribute to academic difficulties and about whether such hypothesized negative effects are likely to be large enough to be educationally meaningful. On the one hand, some have argued that dialect differences are not a major contributor to reading failure and instead point to extrinsic explanatory factors for low reading achievement, including: (a) prejudice against, and lowered educational expectations for, African American students by classroom teachers (Goodman & Buck, 1973; Harber & Beatty, 1978; Ogbu, 1995); (b) inadequate and insensitive instruction (e.g., Cunningham, 1976-1977; Dreeben, 1987; Goodman & Buck, 1973; Sims, 1982); (c) the inappropriateness of testing procedures, and of many tests themselves, for evaluating reading and related abilities of AAVE speakers (Sims, 1982; Smitherman, 1977); and (d) the confounding of socioeconomic and instructional differences with dialect variation in many studies (Dreeben, 1987; Strickland, 1995).

On the other hand, it has also been maintained that a mismatch between the student's and the teacher's or book's representations of linguistic features can indeed make reading acquisition more difficult for an AAVE-speaking student in several ways (Baratz, 1969b; Labov, 1995). Most directly, dialect differences could lead to interference and confusion as children attempt to discover and learn regular spelling–sound correspondences, to identify the oral counterparts of letter strings than can be sounded out, and to comprehend syntactic and semantic relationships in text. Furthermore, in conjunction with other educational barriers that African American children may encounter in classrooms, the mismatches between dialects might also (or instead) have a more indirect effect on the child's motivation and attitude toward literacy, including a loss of confidence in the alphabetic system and reduced educational aspirations (Labov, 1995).

Following several decades of neglect (Rickford & Rickford, 1995), there has recently been a resurgence of interest in the potential educational consequences of dialect differences. In part, this renewed attention to the issue is fueled by increasing concerns about persistent racial and socioeconomic gaps in academic achievement (National Assessment of Educational Progress, 1995, 1997), which are evident from the outset of schooling (Denton & West, 2002; West, Denton, & Germino-Hausken, 2000). As noted in an influential national report (Committee on the Prevention of Reading Difficulties of Young Children, 1998), progress in addressing this achievement gap is still hindered by a lack of strong empirical evidence regarding the relationship between children's usage of AAVE and their acquisition of reading skills, particularly in the early school years. In a thorough and critical review of the pertinent research, Washington and Craig (2001) noted that mixed results have been obtained when researchers have examined the extent to which dialect differences interfere with decoding or reading comprehension. This body of work, it should be noted, consisting of a few dozen studies conducted before 1982, primarily examined samples of older students rather than beginning readers. In light of this dearth of evidence, it has not been possible to draw firm conclusions about the nature and strength of the associations, if any, between AAVE usage and early reading acquisition.

The present study sought to obtain new and stronger evidence for such an association. In designing the research, however, we began with a somewhat different premise from that of prior studies, namely, that although much of value can be learned by examining how consistently a child uses AAVE features in everyday speech, it would also be fruitful to measure the student's familiarity with SE itself, which may be more directly related to success in learning to read. That is, SE is the version of English that will be encountered by the child in most of the written materials used in reading instruction, and many teachers will produce SE forms frequently in their speech when teaching. The clarity of instruction, therefore, is likely to be enhanced for children who already have some familiarity with SE,

and the task of discovering the fundamental correspondences between oral and printed language is likely to be a less daunting one.

We hypothesized, therefore, that some AAVE-speaking children begin school knowing more about aspects of SE that are not mandatory in AAVE and that those children should have an advantage in learning to read. Hence, instead of measuring the frequency with which AAVE features are spontaneously produced by young African American students, we evaluated their familiarity with SE by measuring how readily they could repeat SE sentences verbatim when instructed to do so. There is considerable evidence that imitation of a sentence presented in SE is likely to diverge from the model in speakers who are not fully familiar with SE but who may be able to understand some or all elements of the dialect even if they do not spontaneously produce them (Baratz, 1969a; Labov, Cohen, Robins, & Lewis, 1968; Radloff, 1991). Hence, some aspects of the sentence are likely to be reproduced in a form that is acceptable in the speaker's stronger dialect (in this case, AAVE) but not in SE. The extent to which such dialect differences do not occur during sentence imitation, but rather how often the SE forms are reproduced, was thus chosen as our measure of children's familiarity with SE. Note that the children's actual AAVE usage in less formal situations was not assessed; therefore, it cannot be known whether children whose sentence imitation performance indicates high familiarity with SE might differ from their classmates with regard to the degree to which they produce AAVE features in everyday speech.

Furthermore, consistent with contemporary accounts of the processes and prerequisites of reading acquisition (e.g., Committee on the Prevention of Reading Difficulties of Young Children. 1998; McCardle, Scarborough, & Catts, 2001), we sought to examine children's familiarity with SE separately for its phonological and morphosyntactic features. A beginning reader's phonological skills are known to be critically important for grasping the alphabetic principle (that letters usually represent the phonemes of spoken words) and for discovering and learning the regular spelling–sound relationships that enable printed words to be decoded (National Reading Panel. 2000). However, strong grammatical and vocabulary knowledge also contributes to success in early reading, especially for acquiring strong text comprehension skills (e.g., Catts, Fey, Zhang, & Tomblin, 1999; Scarborough, 1998). Hence, whether a strong relationship between reading and familiarity with SE will be seen might depend on which aspects of them have been measured.

The main goals of this study, therefore, were: (a) to assess familiarity with SE among young African American students from low-SES backgrounds by measuring the degree to which they could reproduce phonological and grammatical features of SE in a sentence imitation context, and (b) to examine whether familiarity with SE, when measured in this manner, is related to reading achievement differences in the early school grades.

Method

Participants

All of the children were African American students in kindergarten through second grade who were attending historically low-performing schools serving low-income communities in three large U.S. cities: Cleveland, Ohio; New Orleans, Louisiana; and Washington, D.C. These school districts were the sites for a school reform initiative directed by the American Federation of Teachers (AFT), a national teacher union. The central aim of that project was to provide professional development in reading instruction to inner-city teachers.

For this study, a random sample was drawn from each kindergarten, Grade 1, and Grade 2 class at two schools in each city, yielding a total sample of 217 children. As shown in Table 1, approximately equal numbers of boys and girls were included at each grade and from each city. High proportions of students in all three cities were eligible for the federal free or reduced lunch program, indicative of the low SES of the communities served by these schools. There was, however, some variation in SES across schools. Nearly all children at both schools in New Orleans were eligible for the lunch program (94% and 100%), but the participating schools were more dissimilar in this respect in Cleveland (84% and 100%) and Washington, D.C. (74% and 94%).

Procedures

All data for this study were collected during April, May, and June of the 2000-2001 academic year when, as part of the AFT project, a variety of measures of reading and related skills was obtained for all children in kindergarten through second grade in the participating schools, although not all children were given every measure. These assessments were undertaken primarily for program evaluation purposes and were conducted by 11 experienced reading teachers (8 African American, 3 White) who had been released from regular classroom responsibilities

Table 1
Sample Description

		Cleveland	New Orleans	Washington	Total
% Free/reduced lunch		92.0	96.8	82.3	91.1
Kindergarten	Boys	19	13	8	40
	Girls	17	12	9	38
	Total	36	25	17	78
First grade	Boys	11	11	12	34
	Girls	18	9	11	38
	Total	29	20	23	72
Second grade	Boys	14	6	14	34
	Girls	15	10	8	33
	Total	29	16	22	67
Total	Boys	44	30	34	108
	Girls	50	31	28	109
	Total	94	61	62	217

to participate in the professional development project in their districts. All tests were individually administered in one or two brief sessions per child (15–30 min) during school hours in the libraries or other quiet areas of the children's schools.

Measures

Reading achievement Three subtests of the Woodcock Reading Mastery Tests–Revised (WRMT–R) were used to assess how well the children could read aloud printed words of increasing difficulty (Word Identification), phonologically decode printed pseudowords (Word Attack), and indicate their comprehension of short text passages by providing spoken words that would be semantically and syntactically appropriate at locations marked by blanks in the printed texts (Passage Comprehension). The WRMT–R is a widely used, nationally standardized test for which strong reliability and validity have been demonstrated (Woodcock, 1987). Dialect-sensitive scoring of the Passage Comprehension items was used, as recommended in the test manual. Raw scores were converted to Rasch-scaled (equal-interval) *W* scores for analysis. The Passage Comprehension test was not administered to the kindergartners.

Sentence imitation. Fifteen sentences to be imitated were presented in a picture book context, using a spiral-bound book constructed for this purpose. The sentences, listed in Table 2, ranged in length from 3 to 16 words ($M = 9.3$) and were constructed to include many phonological and morphosyntactic elements that are often produced differently in AAVE and SE. To minimize vocabulary effects in this task, we selected words that in our judgment would be familiar to all children at this age. The median frequency for the 52 open-class words was 239 per million words of text (Carroll, Davies, & Richman, 1971); the words *raisin* (3 per million), *snack* (5), and *bathroom* (7) were least frequent but were deemed familiar to most children even though they do not often occur in written texts.

The seven book illustrations showed: (a) a boy and a girl standing beside a bicycle (Sentences 1, 2, 3); (b) the children riding bicycles, with the boy leading (Sentences 4, 5); (c) the children washing up at a sink (Sentences 6, 7); (d) the children preparing snacks at a table (Sentences 8, 9, 10); (e) the girl peering into a refrigerator, contents not visible (Sentences 11, 12); (f) the children seated at the table using raisins to decorate open-faced peanut butter sandwiches (Sentences 13, 14); and (g) the two sandwiches, with raisins arranged as a flower on one and as an elephant on the other (Sentence 15). The two children in the pictures are of indeterminate ethnicity, and their clothing, activities, and locations were selected and drawn to be equally familiar to all children regardless of geographic region or SES.

The child was asked to repeat each sentence immediately after the examiner presented it, saying the sentence exactly the way the examiner had. Two or more practice sentences were given to clarify the instruction. A demonstration tape was used to train examiners in all cities to present the sentences in a uniform way. They were instructed to speak slowly with moderate expression and with overly precise articulation of all words. Of the eight examiners who administered the task, five were African American and three were White.

From audio recordings of the administration of this measure, all occurrences of verbatim reproductions, dialect differences, and memory errors in the children's imitations were coded. Two scorers (one African American, one White) independently coded 40% of the sessions so that interscorer reliability could be evaluated. From the large initial pool of potential dialect difference items (underlined in Table 2), many were then eliminated for various reasons. First, several had unacceptable rates of examiner error in presenting the sentences (e.g., regardless of race, testers rarely articulated the final /t/ of *must* in Sentence 7). Second, interscorer reliability was difficult to achieve in some cases, and it was decided to eliminate all items for which there was less than 80% agreement between scorers as to whether the child had produced a dialect difference or a verbatim reproduction in recalling the item. Third, for some items many children did not recall the portion

Table 2
Sentence Imitation Items Scored for Dialect Differences and Memory Errors

The 15 sentences to be imitated, with original pool of phonological (P) and grammatical (G) items underlined, and the 60 retained[a] items in boldface

1. This **is** Joe.
   1a
2. The girl behind him **is call - ed** Lisa.
   P **2a** **2b** **2c** **2d**
3. She is **Joe's** best friend.
   G **3a** **3b** P
4. Joe **ride - s his** bike down **the street** really **fast.**
   4a **4b** **4c** **4d** **4e** G **4f**
5. Lisa pushes hard **because** she is trying to **keep** up **with** Joe.
   G P **5a** G **5b** **5c**
6. Both of the **kid - s are very** hungry, so **they are** going to make **them-selves** a snack.
   P P P **6a** G **6b** **6c** **6d** **6e** P **6f** **6g**
7. First, they must wash **their hand-s** in **the bath**-room.
   P P **7a** **7b** **7c** **7d**
8. In the **kitchen**, Lisa spread - **s** peanut butter on **two** slice - **s of** bread.
   P **8a** P **8b** **8c** G **8d**
9. Joe **pour-s him-self** some milk **with**-out spilling any.
   9a **9b** **9c** **9d** G
10. He **pour-ed another** glass for Lisa.
    10a **10b** P
11. **Then** Joe **ask** - **ed**, "**Isn't there** any jelly?"
    11a **11b** **11c** **11d** **11e** G
12. **Lisa** answer-ed, "**We** don't have **any** jelly **so let's** have raisins instead.
    12a G **12b** P,G **12c** **12d** **12e** G
13. Lisa used **raisin-s** to draw a flower on **her peanut butter.**
    G **13a** P **13b** **13c**
14. Joe **decide** - d to make **an** elephant with an **open** mouth and strong **leg** - s.
    14a G P **14b** P G **14c** P P **14d** G
15. Joe think - **s** that the **snack-s are now** ready **to** eat.
    P **15a** P P **15b** **15c** **15d** **15e**

The 21 phonological items retained for analysis	
4d, 5c, 6d, 7c, 7d, 9d, 10b, 11e	Substitution of /t/, /d/, /f/, /v/, /s/, or /z/ for *th* (/θ/ or /ð/).
2a, 3b, 4e, 4f, 11b, 11d, 12e	Omission or reduction of consonant cluster.
1a, 4a, 4c, 8d, 13b, 15e	Omission of singleton consonant.

The 22 grammatical items retained for analysis	
2b, 6b, 6e, 15c	Omission or substitution of copula or *be* auxiliary verb.
2d, 4b, 8b, 9a, 10a, 11c, 15a	Omission of verb tense inflection (*-ed* or *-s*).
3a, 7b, 13a, 15b	Omission of possessive or plural inflection.
6f, 6g, 7a, 9c, 11e	Omission of reflexive, or change of pronoun.[b]
12c	Substitution of *none* or *no* for *any.*
14b	Substitution of *a* for *an*.

The 17 memory items (omission or substitution of the word)
2c, 5a, 5b, 6a, 6c, 8a, 8c, 9b, 11a, 12a, 12b, 12d, 13c, 14a, 14c, 14d, 15d

[a]Of these, three phonological (1a, 4a, 11b) and three grammatical (6f, 11c, and 14b) items did not correlate well with total scores or with other items and thus were dropped.
[b]Both "Isn't they any jelly?" and "Isn't it any jelly?" were often produced when "Isn't there any jelly?" was imitated (Item 11e).

of the sentence in which the targeted phonological or morphosyntactic form was embedded (e.g., the phrase *In the kitchen* was often not recalled; therefore, phonological differences in the articulation of *the* in that phrase could be scored for few children). Fourth, more than enough exemplars of some types of items remained (e.g., plurals); therefore, only a subset of these was retained in the final item set.

After these steps were taken, a total of 21 phonological items and 22 grammatical items remained for analysis. Next, 17 memory items were selected on the basis of high interscorer reliability and a suitable range of response accuracy in the sample. Strong reliability (percentage agreement) was obtained for all three sets of items: $M = 92.2\%$ (range = 82–100) for the 21 phonological forms, $M = 94.8\%$ (range = 86–100) for the 22 grammatical forms, and $M = 98.6\%$ (range = 96–100) for the 17 memory items.

Three summary scores were then computed. Each represented the percentage of times a child produced a particular kind of response out of the total number of opportunities to do so. To compute the phonological score, only the items on which the child actually recalled the item (word or morpheme) in some form (verbatim or altered) were included. The number of verbatim reproductions (rather than phonological dialect differences) the child produced was divided by the number of eligible items (maximum 21), and this ratio was multiplied by 100. For the grammatical score, the computation was analogous except that nonrecalled items were included in the eligible set if: (a) the omission represented a known feature of AAVE (e.g., omission of a copula or reflexive), and (b) the surrounding portion of the sentence was imitated. For the memory score, the number of items correctly recalled was divided by the number of memory items; all 17 words in the set were included unless the entire clause or sentence containing an item had not been imitated, as occasionally occurred.

Story recall. At the conclusion of the sentence imitation task, the child was asked to tell "everything you remember about what happened" in the story, using the picture book to guide recall. Examiners were permitted to provide up to three "Anything else?" prompts to reticent children during story recall. Using a checklist, the examiner recorded which of 32 "gist" recall elements were included, at any point, during the child's narrative. These were: use of the name *Joe*; use of the name *Lisa*; she is behind him; they are good/best friends; he goes fast; she pushes/tries hard; so (causal link); she wants to keep up with him; they are hungry; so (causal link); they make/have a snack; first (temporal); must/gotta (necessity); wash their hands; in the bathroom; in the kitchen; she puts peanut butter on bread; he pours/gets milk; for himself; for her; doesn't spill; he asks for jelly; she answers; she says there's no jelly; she suggests raisins instead; she draws with raisins; she makes a flower; he makes an elephant; with a big mouth; with strong/big legs; he thinks/says; it's time to eat. By design, the illustrations provided little of this information, all of which was conveyed verbally in the 15 sentences. Hence, relying on the pictures alone would enable a child to earn few points on the recall task. This story recall measure was included so that we could investigate: (a) whether the production of dialect differences during sentence imitation was associated with comprehension of the story, and (b) whether the hypothesized associations of dialect differences with reading might simply reflect a more general tendency that also extended to nonprint verbal abilities.

Results

The analyses proceeded in four phases. First, reading achievement scores were examined for possible city and school differences to determine whether these factors might need to be controlled in subsequent analyses. Second, correlations among individual items in the phonological and grammatical imitation measures were computed so that inappropriate items could be identified and dropped. Third, sentence imitation scores were analyzed for effects of grade, city, school, SES, and examiner differences. Last, the relationships of familiarity with SE to reading achievement and story recall were examined. Before analysis, distributions of all measures were examined for the presence of severe skewness, outliers, and other distributional irregularities, and no such threats to validity were observed.

Reading Achievement

Table 3 provides a summary of reading scores at each grade in each city. At each grade, the available reading measures (Word Identification, Word Attack, and except in kindergarten, Passage Comprehension) were entered into a one-way multivariate analysis of variance (MANOVA) with city as the between-group factor. City difference were negligible in Grade 1, $\eta^2 = .03$, Wilks's $\lambda = .94$, $F(6, 270) = 1.42$, $p = .21$, and in Grade 2, $\eta^2 = .04$, Wilks's $\lambda = .96$, $F(6, 116) = 0.43$, $p = .86$. In kindergarten, however, reading scores were not equivalent across cities, $\eta^2 = .16$, Wilks's $\lambda = .71$, $F(4, 144) = 6.71$, $p < .001$. These differences occurred for both Word Identification and Word Attack, and post hoc Tukey tests indicated that scores were higher in Cleveland than elsewhere.

Correlations of Sentence Imitation Scores With Performance on Individual Items

Because only recalled items were included when calculating the phonological and grammatical scores

Table 3
Mean Scores on End-of-Year Woodcock Reading Mastery Tests–Revised Reading Tests

		Word Identification		Word Attack		Passage Comprehension	
Grade	City	*W*	GE	*W*	GE	*W*	GE
K	CL	402.3 (27.1)	1.4	458.9 (16.2)	1.2		
	NO	379.1 (24.5)	1.0	443.5 (12.2)	K.7		
	DC	371.1 (21.8)	K.9	446.5 (14.8)	K.8		
1st	CL	432.8 (24.2)	1.9	472.3 (16.8)	1.7	453.9 (15.8)	1.6
	NO	419.4 (23.8)	1.7	464.3 (21.0)	1.4	450.2 (17.7)	1.5
	DC	430.5 (27.7)	1.9	472.1 (19.8)	1.7	460.8 (15.2)	1.8
2nd	CL	448.6 (19.2)	2.4	475.7 (18.1)	1.8	469.0 (12.3)	2.3
	NO	456.0 (17.7)	2.6	480.3 (16.4)	2.0	469.1 (10.3)	2.3
	DC	452.7 (26.6)	2.5	478.2 (21.7)	1.9	469.4 (16.5)	2.3

Note. K = kindergarten; CL = Cleveland, Ohio; NO = New Orleans, Louisiana; DC = Washington, D.C.; *W* = Rasch-scaled scores; GE = grade equivalent associated with the mean *W* score. Standard deviations are shown in parentheses.

on the sentence imitation task, these percentages were often based on slightly different item sets for different children. Hence, conventional measures of internal consistency could not be applied. Instead, point–biserial correlations were first computed between each item and the total phonological and grammatical scores. Then, for each pairing of items, a phi coefficient was obtained, indicating whether children who produced a dialect difference when recalling a particular item also tended to produce a dialect difference on another item.

Three of the 21 phonological items (1b, 4b, and 11c) were found to correlate weakly with the summary score, $rs(215) = .21$, .23, and .15, respectively. Consistent with this, only one phi coefficient larger than .25 was obtained between these items and any of the others. It was therefore decided to drop those three items from the scale, retaining the other 18 items for which higher correlations with total phonological scores were obtained, $rs(215) = .30$ to .61, $p < .001$. Similarly, 3 of the 22 grammatical items (6f, 11c, and 14b) were not well correlated with the summary score, $rs(215) = .04$, .27, and .15, respectively, and were only weakly related to other items. Hence, these were dropped, leaving 19 grammatical items for which stronger associations with total grammatical scores were seen, $rs(215) = .31$ to .64, $p < .001$.

In refining these measures, we were also concerned that many grammatical dialect differences also involve phonological differences (Owens, 1996). For example, omitting a plural inflection can also be construed as the omission of a final consonant or the reduction of final consonant cluster (Labov, 1972). We therefore sought to determine whether responses for any items assigned to the grammatical set actually covaried more closely with responses on phonological items. To that end, correlations between total phonological scores and individual morphosyntactic items were computed. With one exception, every such item correlated more highly with the total grammatical score (median $r = .47$) than with the total phonological score (median $r = .30$). The exception was Item 2d, on which the dialect difference involved omission of the past tense *-ed* affix; this item correlated well with both the grammatical (median $r = .47$) and phonological (median $r = .49$) scores. Because other grammatical items of this sort did not show the same pattern, however, it was decided not to reassign this item.

For each child, therefore, a revised phonological score, based on 18 items, and a new grammatical score, based on 19 items, were computed. These measures were well correlated with each other but far from perfectly, in the entire sample, $r(215) = .66$, and within each age group: $r(76) = .57$ for the kindergartners, $r(70) = .68$ for the first graders, and $r(65) = .59$ for the second graders. These two scores were thus examined separately in all subsequent analyses.

The memory score was found to have adequate internal consistency (Cronbach's $\alpha = .74$) and was correlated ($p < .001$) with both the phonological (median $r = .31$) and the grammatical (median $r = .50$) dialect difference scores that were derived from the same sentence imitation task.

Sentence Imitation: Demographic Differences

In the sample as a whole, phonological and grammatical scores each ranged from 7% to 100%, but more imitations of SE forms (and hence fewer

productions of dialect differences) occurred on grammatical ($M = 61.6$, $SD = 22.6$) than phonological ($M = 50.8$, $SD = 19.9$) items. Verbatim imitation of 75% or more of the items, indicating strong familiarity with SE, was seen for 34% of the children on the grammatical scale and for 11% on the phonological scale. At the other extreme, dialect differences were produced at least 75% of the time by 7% of the children on the grammatical scale and by 11% on the phonological scale.

Table 4 provides a summary of sentence imitation performance by students in each grade from the three cities. Each score was entered into a separate 3 (grade) × 3 (city) ANOVA, with Tukey tests used for follow-up comparisons. Main effects of grade were found for the phonological score, $\eta^2 = .105$, $F(2, 208) = 12.26$, $p < .001$, and grammatical score, $\eta^2 = .211$, $F(2, 208) = 27.82$, $p < .001$. For each measure, means were lower in kindergarten than in the other two grades, in which scores were similar. Differences among the three cities were also found: for phonological scores, $\eta^2 = .274$, $F(2, 208) = 39.17$, $p < .001$, and for grammatical scores, $\eta^2 = .061$, $F(2, 208) = 6.73$, $p = .001$. In neither analysis was there a significant interaction of grade with city ($\eta^2 < .015$).

Because local teachers did the testing in each district, the observed variation between cities could have been attributable to examiner differences. By happenstance, when an examiner in New Orleans fell ill, one from Cleveland traveled to New Orleans and assisted with the assessments there for a few days. It was therefore possible to compare scores for the children that she tested in each location. She administered the sentence imitation measure to only one New Orleans kindergartner, however; therefore, these analyses included only first and second graders. Mean phonological scores for Cleveland and New Orleans, respectively, were: 62.9 ($SD = 17.8$, $N = 24$) versus 40.1 ($SD = 16.9$, $N = 6$) for Grade 1, and 61.3 ($SD = 16.0$, $N = 21$) versus 49.1 ($SD = 16.5$, $N = 10$) for Grade 2. For the grammatical score, the corresponding values were: 73.1 ($SD = 18.0$) versus 57.6 ($SD = 22.8$) and 74.3 ($SD = 18.1$) versus 62.0 ($SD = 24.2$). When these data were entered into 2 (city) × 2 (grade) ANOVAs, city differences were found for both phonological scores, $\eta^2 = .175$, $F(1, 57) = 12.08$, $p < .001$, and grammatical scores, $\eta^2 = .090$, $F(1, 57) = 5.67$, $p = .021$. These effects were similar in magnitude to those seen for the entire sample, suggesting that examiner differences did not account for the finding that children from New Orleans exhibited less familiarity with SE on our measures than did students from the other cities from which the samples were drawn.

Another concern was that the race of the examiner could affect the extent to which a child might produce dialect differences during sentence imitation. This issue was addressed by examining scores for Cleveland and New Orleans, in which both African American and White examiners had administered the measure (in Washington, D.C., all examiners were African American). No effect of tester's race was obtained for either phonological scores, $\eta^2 = .005$, $F(1, 146) = 0.66$, $p = .42$, or grammatical scores, $\eta^2 = .002$, $F(1, 146) = 0.25$, $p = .62$.

Table 4

Percentage of Phonological and Grammatical Items on Which the School English Form Was Reproduced, and Percentage Correct on Memory Items, in Children's Imitations of Sentences Presented in School English

		Phonological (18 items)			Grammatical (19 items)			Memory (17 items)		
Grade	City	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
K	CL	53.2	17.4	21–94	51.8	22.2	18–95	58.1	17.5	24–93
	NO	27.8	10.8	11–55	41.5	16.8	20–74	46.9	19.5	6–81
	DC	43.6	18.4	25–100	48.6	23.7	7–100	60.0	16.3	41–94
	Total	43.0	19.2	11–100	47.8	21.2	7–100	54.9	18.6	6–94
1st	CL	62.2	17.9	23–94	70.7	18.7	29–100	64.9	18.2	25–87
	NO	36.8	17.0	13–65	56.3	23.7	18–100	58.6	21.4	25–100
	DC	59.6	20.4	25–100	72.6	18.7	19–94	63.9	20.8	29–100
	Total	53.3	21.3	13–100	67.3	21.1	18–100	62.8	19.9	25–100
2nd	CL	64.2	14.8	35–94	73.6	16.8	44–100	73.2	17.1	37–94
	NO	43.2	16.0	17–78	64.7	19.9	23–89	66.0	19.3	24–100
	DC	54.4	12.1	31–78	73.9	15.9	42–95	73.0	18.2	29–93
	Total	56.0	16.3	17–94	71.6	17.5	23–100	71.4	18.0	24–100

Note. K = kindergarten; CL = Cleveland, Ohio; NO = New Orleans, Louisiana; DC = Washington, D.C.

For memory scores, there were differences among grades, $\eta^2 = .103$, $F(2, 208) = 11.93$, $p < .001$, and cities, $\eta^2 = .04$, $F(2, 208) = 4.29$, $p = .015$, but no interaction between the factors, $\eta^2 = .005$, $F(2, 208) = 0.26$, $p = .90$. As one would expect, each grade differed from the others, with error rates highest for kindergartners and lowest for second graders. Post hoc tests also indicated that at all grades, memory scores were somewhat lower in New Orleans than in the other districts.

Story Recall

The mean number of gist items that were included in the child's retelling of the story was lower in kindergarten ($M = 16.5$, $SD = 6.0$) than in first grade ($M = 19.3$, $SD = 5.4$) and second grade ($M = 19.9$, $SD = 5.2$), $\eta^2 = .115$, $F(2, 208) = 13.39$, $p < .001$. Mean scores in Cleveland were 3 to 4 points higher than those in the other two cities, $\eta^2 = .125$, $F(2, 208) = 14.89$, $p < .001$, but there was no significant Grade × City interaction, $\eta^2 = .031$.

SES Differences

Within the districts in which the two schools were not equal in SES (i.e., the proportion of students eligible for the federal lunch program), mean reading scores at those schools were nearly identical on all reading tests, regardless of grade (all η^2s $< .003$, $F < 0.26$, $p > .61$). Likewise, neither story recall nor memory scores showed any SES differences or interactions with grade when schools in these cities were compared (all η^2s $< .01$, $F < 0.8$, $p > .39$).

In contrast, we did obtain SES differences between these schools for both phonological scores (60.6 vs. 53.2), $\eta^2 = .04$, $F(1, 145) = 6.76$, $p = .01$, and grammatical scores (69.6 vs. 60.5), $\eta^2 = .05$, $F(1, 145) = 7.66$, $p = .006$, from the sentence imitation task. Interactions of SES with grade were negligible and nonsignificant, all η^2s $< .015$, $F(2, 145) < 1.1$, $p > .35$.

Relationships of Familiarity With SE to Reading Achievement and Story Recall

Table 5 contains three correlation matrices that show the relationships of reading and story recall measures to our measures of familiarity with SE at each grade. First, the bivariate (zero-order) correlations are shown. In kindergarten and first grade, the phonological and grammatical sentence imitation scores were each correlated, to about the same degree, with all three measures of reading ($rs = .42$ to .59). In second grade, however, all reading scores were again related to grammatical scores ($rs = .34$ to .49), but effects were weaker and, except for Passage Comprehension, nonsignificant for phonological scores ($rs = .08$ to .29). Dialect difference scores were not strongly related to story recall ($rs = .07$ to .28) at any age.

Second, because demographic differences were observed in some prior analyses, partial correlations were computed that controlled for city and SES of the school. As shown in the second matrix in Table 5, the pattern of results was generally similar.

Third, as one would expect, memory scores from the sentence imitation task were also correlated with the reading measures for the kindergartners ($rs = .28$ to .32, $p < .05$) and first graders ($rs = .26$ to .30), although this was not seen in the second grade sample ($rs = .01$ to .21). Therefore, to demonstrate that the observed relationships of reading to grammatical and phonological scores do not simply reflect a general verbal imitation ability, the last matrix in Table 5 shows correlations from which memory scores as well as city and school differences have been partialed out. As can be seen, the pattern of results was not substantially altered.

As one would expect in the early grades, scores on the reading subtests were not independent. Word Identification and Word Attack were strongly correlated ($rs = .79$, .85, and .73 for kindergarten, first grade, and second grade, respectively), and Passage Comprehension was related to both Word Identification and Word Attack in Grade 1 ($rs = .90$ and .85, respectively) and in Grade 2 ($rs = .76$ and .73, respectively). In first grade, after the two other reading scores had been entered into a regression analysis, almost no additional variance (less than 1%) in Passage Comprehension was accounted for by either measure of familiarity with SE from the sentence imitation test, indicating that word recognition skills were the primary limiting factor on reading comprehension at this age. In second grade, however, an additional 2.8% of variance was accounted for by phonological scores, $F(1, 59) = 4.95$, $p = .03$, and an additional 4.4% by grammatical scores, $F(1, 59) = 8.22$, $p = .006$, above and beyond the contribution (64.2%) of Word Identification and Word Attack.

Discussion

We found wide variation among young, urban African American children from low-income families in the degree to which they produced SE rather than AAVE forms when imitating sentences presented in SE, indicating that they differed substantially in their familiarity with SE during the early school years. As

Table 5
Correlations of Sentence Imitation Scores With Reading Skills and Story Recall

Score	Grade	Word Identification	Word Attack	Passage Comprehension	Story Recall
		Zero-order correlations			
Phonological	K	.44***	.49**		.25*
	1st	.45***	.42**	.48***	.07
	2nd	.08	.22	.29*	.20
Grammatical	K	.42***	.51***		.20
	1st	.58***	.50***	.59***	.28*
	2nd	.34**	.36**	.49***	.28*
		Partial correlations controlling for city and school SES differences			
Phonological	K	.39**	.41**		.04
	1st	.44***	.41**	.50***	.03
	2nd	.16	.32*	.34**	.19
Grammatical	K	.47***	.52***	–	.16
	1st	.61***	.51***	.59***	.27*
	2nd	.38**	.39**	.52***	.29*
		Partial correlations controlling for city, school SES, and memory score differences			
Phonological	K	.35**	.38**		.09
	1st	.42**	.38**	.47***	.09
	2nd	.16	.31*	.32*	.16
Grammatical	K	.38**	.49***		.11
	1st	.58***	.46***	.53***	.18
	2nd	.41**	.36**	.48***	.17

Note. K = kindergarten; SES = socioeconomic status.
*$p < .05$, two-tailed. **$p < .01$, two-tailed. ***$p < .001$, two-tailed.

hypothesized, reading achievement was well correlated with children's familiarity with SE, and these relationships could not be attributed simply to differences in children's memory abilities. The results are consistent with the long-standing view that learning to read may indeed be more difficult for this population of students, and they raise questions about the sources of variation in familiarity with SE and about mechanism(s) by which reading acquisition may be impeded for those children who are less familiar with SE when they enter school.

Variation in Children's Familiarity With SE

Some of the children in each grade were able to imitate SE forms a high proportion of the time. Others, however, frequently produced an AAVE form in lieu of the SE model, and a majority produced dialect differences and verbatim responses in roughly equal proportions. In other words, even in their first few years of school, most of these children were conversant with some features of SE, though few were fully able to imitate all of the phonological and morphosyntactic forms that we examined.

Traditionally, social factors, especially SES and exposure to speakers of other dialects, have been linked to dialect variation among individuals (Baugh, 1983; Dillard, 1972; Labov & Harris, 1986; Sims, 1982; Washington & Craig, 1998; Washington et al., 1998; Wolfram & Fasold, 1974). Our finding that children at the schools with the highest proportions of students in the federal lunch program had lower phonological and grammatical scores indicates that familiarity with SE (but neither memory nor reading skill) was indeed related to SES, even over the relatively narrow range in our sample.

Additional evidence for socioeconomic differences in children's level of familiarity with SE can be found in some preliminary data from an ongoing study in which our sentence imitation task was administered to first graders attending suburban schools in lower-middle-class communities in Connecticut (B. Swainson, personal communication, December 2, 2003). In an African American sample ($N = 48$) from a school with primarily minority students and at which 27% of the enrollment qualified for the federally lunch program, mean scores were higher and variances a bit lower than those for first graders in our less advantaged urban sample: 85.1

($SD = 17.1$, range = 24–100) versus 67.3 ($SD = 21.1$, range = 18–100), respectively, for the grammatical scale; 65.0 ($SD = 16.2$, range = 30–100) versus 53.3 ($SD = 21.3$, range = 13–100), respectively, for the phonological scale. These comparisons suggest that familiarity with SE increases with SES over a broader range than could be observed within our sample but that wide individual differences exist even among African American children from more middle-class suburban communities.

Moreover, in Swainson's White sample ($N = 25$), drawn from a school with few minorities but similar SES (26% eligible for the federal lunch program), means were very high, and variability was much reduced, for both grammatical ($M = 96.4$, $SD = 3.9$, range = 88–100) and phonological ($M = 88.4$, $SD = 7.5$, range = 73–100) sentence imitation scores. Memory scores of the White children ($M = 74.7$, $SD = 17.1$) were similar in level and variability, however, to those for her African American sample ($M = 76.4$, $SD = 15.1$), indicating that the groups were of similar general ability. These data suggest that racial as well as SES differences matter, such that growing up in an African American family or in a community with a predominantly African American population may provide fewer opportunities to gain familiarity with SE through listening and speaking in the child's early years. Different language norms and expectations, including whether the child is expected to use SE in school, may also be conveyed to young African American children from different socioeconomic strata.

The geographic differences that we observed can probably be explained in terms of exposure to dialect variation and language norms in the community, although regional childrearing differences cannot be ruled out. In New Orleans, fewer children in each grade reproduced large numbers of the SE forms when imitating the phonological and morphosyntactic items that we scored. It is well established that more so than in other regions, many Whites in New Orleans and other southern U.S. cities speak a dialect that shares many features with AAVE (Wolfram, 1974). It is highly probable, therefore, that fewer models of SE may have been available in the school and the community to the children from New Orleans than to those in Cleveland and Washington, D.C.

Much variance in our sample remains unexplained by SES, race, and regional differences, however. There is evidence that even within restricted socioeconomic strata, variation in parental characteristics and behaviors is related to the growth of children's language and other aspects of school readiness (e.g., Hart & Risley, 1995), although this relationship is generally a modest one (Scarborough & Dobrich, 1994). We would thus expect that the educational levels, childrearing practices, and language patterns of children's parents, which were not measured in this study, would account for some additional variance in children's familiarity with SE.

The degree of variability (i.e., ranges and standard deviations) for both phonological and grammatical scores was similar in kindergarten, first grade, and second grade, suggesting that the experience of schooling does not markedly reduce the wide variability among children in their familiarity with SE. Mean scores, however, did show cross-sectional differences between the end of kindergarten and the end of first grade (53% vs. 43%, respectively, for phonological scores, and 56% vs. 39% for grammatical scores) but not over the following year. These results suggest that familiarity with both phonological and grammatical aspects of SE increases, although not dramatically, in conjunction with the onset of schooling.

We originally hypothesized that developmental changes would be driven by exposure to SE in the teachers' speech within classrooms. There are reasons, however, to doubt this. Although we did not directly measure the language usage of the teachers in the participating schools, firsthand observations were provided by AFT personnel who visited these sites on many occasions. In Cleveland, it was reported, nearly all primary grade teachers were White and their speech included very few AAVE features. In contrast, all teachers in Washington, D.C. were African American, and most were observed to produce many features of AAVE as well as of SE in their speech in the classroom. (In New Orleans, the race and language patterns of teachers were more mixed than in the other two cities.) The children in Washington, D.C. were thus presumably exposed to many fewer models of SE in their classrooms than were the children from Cleveland. Despite the starkly different situations in these cities, there was no City × Grade interaction in our analyses. In fact, somewhat larger differences between kindergarten and first grade means were seen in Washington, D.C. than in Cleveland (15.6 vs. 7.5 percentage points for phonological scores, and 19.1 vs. 15.4 for grammatical scores, respectively).

What all of the classrooms had in common, however, because of their participation in the professional development program, was the availability and use of books and other print materials for reading instruction. It thus seems likely that exposure to print itself, and gains in knowledge about print and reading, may be more responsible for

increasing familiarity with SE in the early school years than is exposure to oral models. If so, however, it is puzzling that no further increases were seen between first and second grades, despite continued print exposure during that year. From the available evidence, therefore, no definitive conclusions can be drawn about this hypothesis or others concerning the development of dialect knowledge in the early school years. Because the design of this study was cross-sectional, longitudinal comparisons are needed to explore this issue more thoroughly. Moreover, we did not examine changes in dialect differences from the very start of schooling, and one might expect that marked gains in familiarity with SE might occur during the kindergarten year.

Finally, because we did not measure children's production of AAVE features in their colloquial language (e.g., in conversation with peers), we do not know how AAVE usage might be related to variation in familiarity with SE as measured in our sentence imitation task. Whether children who can imitate many SE forms also produce fewer AAVE features in their everyday speech remains an open question.

Associations Between Reading Achievement and Familiarity With SE

As hypothesized, there were consistent and substantial correlations between early reading skills and the production of dialect differences during sentence imitation, our measure of familiarity with SE. The magnitudes of these associations, especially in kindergarten (mean $r = .465$) and first grade (mean $r = .503$), are worthy of comment because they are as large as, or exceed, the effects that are typically obtained at these ages for more conventional cognitive and linguistic predictors of achievement. For example, in a recent meta-analysis (Swanson, Trainin, Necoechea, & Hammill, 2003), average correlations with reading scores were .44 to .48 for phonological awareness, .44 to .55 for rapid naming, .32 to .41 for vocabulary, and .42 to .48 for memory span. Hence, the expected relationship of familiarity with SE to early reading achievement was not only confirmed, but perhaps more important, was found to be large and educationally meaningful by conventional standards.

Although substantial, the findings are correlational and thus do not, in and of themselves, cast light on the nature of the relationship between early reading and dialect knowledge. Many plausible explanations for the observed association can be advanced, including some that are consistent with the findings and some that account less well for the patterns we observed.

The results make it clear, first, that familiarity with SE varies across individuals, not just between racial or socioeconomic groups. Thus, poverty and dialect differences appear to be separable factors, each related to reading achievement but in a different way. It is well established that the relationship between SES and academic achievement is primarily a between-school phenomenon (Committee on the Prevention of Reading Difficulties in Young Children, 1998; White, 1982) because correlations between SES and achievement are much stronger when the school is the unit of analysis ($r = .68$) than when variation among individual students is analyzed ($rs = .22$ to .31). In contrast, our results reflect wide individual differences among individual students from a narrow SES range. These findings thus suggest that although group differences in reading achievement are associated with racial and SES differences (National Assessment of Educational Progress, 1995, 1997), individual variation in reading achievement depends on other factors as well. Among African American students from low-income communities, familiarity with SE may be one such student characteristic. Although no firm conclusion about causality can be drawn from the correlational evidence at hand, several hypotheses can be considered as potential explanations for the observed relationships of dialect differences to reading achievement in our sample.

Global ability differences. The results are not consistent with the hypothesis that general verbal or memory ability underlies the observed correlations between reading and familiarity with SE. First, the vocabulary level of the sentence imitation task was deliberately chosen so as not to be a challenge for the participants, making it unlikely that phonological and grammatical scores would be affected by differences in word knowledge. Second, recall of the gist of the story was not strongly related to measures of familiarity with SE; therefore, the relationship of phonological and grammatical scores to reading apparently did not arise simply because children who understood the sentences better were able to reproduce SE forms more often. Third, the strength of the correlational results was not much affected even when memory scores (based on the recall of particular dialect-neutral words from the imitated sentences) were statistically controlled in the analyses. In short, the observed relationships between familiarity with SE and reading apparently do not simply reflect a general tendency for better students to do well on all verbal tasks.

Instructional quality and teacher bias. Differences in the quality or quantity of reading instruction are probably another major factor in explaining achievement differences between groups and may be the most potent explanation for the worrisome achievement gaps (e.g., Dreeben, 1987) between racial and SES populations. With regard to individual differences too, instructional factors may be important, such that even within a school or a single classroom the quality or quantity of reading instruction may vary across individual students. This may account not only for some of the variability we observed in children's reading scores in our samples but also for the associations between reading and familiarity with SE. In particular, it has often been hypothesized that such a relationship could be a result of teacher bias against AAVE-speaking students in favor of classmates who exhibit greater knowledge and use of standard grammatical and phonological forms (e.g., Cross, DeVaney, & Jones, 2003; Goodman & Buck, 1973; Harber & Beatty, 1978; Ogbu, 1995).

There is abundant evidence that people (including teachers and teachers in training) readily make attributions about a speaker's intelligence, education, and other personal characteristics solely on the basis of listening to a brief excerpt of the individual's speech or oral reading (Cross et al., 2003; Tucker & Lambert, 1969). (Although studies of this sort have not obtained judgments about child speakers, it is likely that listener attributions would be similar to those obtained for adult speakers.) Consistently in this research, speakers of SE have been rated more favorably with regard to their cognitive abilities and social status than have speakers of AAVE or other nonstandard dialects. Moreover, this preference for the standard dialect is exhibited to about the same degree by both White and African American listeners. (Judgments about other traits, however, such as honesty and friendliness, tend to be more influenced by the listener's race.) Hence, in our sample, we would expect that linguistic bias in favor of SE would occur about as often among White as among African American teachers. It is thus not inconsistent with the bias hypothesis that reading achievement scores, and their correlations with measures of familiarity with SE, were similar for the students of the predominantly White teachers in Cleveland and of the African American teachers in Washington, D.C.

Linguistic bias by teachers has been hypothesized to result in a variety of instructional consequences. Because biased teachers are likely to attribute lower capability to AAVE-speaking students and to have negative expectations for their educational success (Baugh, 1999; Cazden, 1988; Taylor, 1983), less instruction may be provided to these students, setting in motion a self-fulfilling prophecy. The quality as well as quantity of reading instruction may also be lower than what the teacher provides to students who are seen as more promising simply because they are more familiar with SE. Particularly detrimental to reading acquisition may be a failure by the teacher to distinguish dialect differences from true reading errors in children's oral reading (Cunningham, 1976-1977; Goodman & Buck, 1973; Meir, 1999; Sims, 1982), a technique that is thought to impede reading acquisition and not to increase knowledge of SE (Delpit, 1998; Piestrup, 1973). When a linguistically biased teacher deprives AAVE-speaking children of sufficient and appropriate reading instruction because they are perceived as less able to benefit from it, those children will be less likely to be successful learners, bringing about a correlation between reading achievement and dialect differences. The findings of our study are consistent with this prediction. Direct confirmation of the bias hypothesis, however, would also require the measurement of teacher attitudes and their differential instructional consequences for children with greater or lesser familiarity with SE.

Linguistic interference between AAVE and SE. The hypothesis that originally motivated this research was that children who are more familiar with SE forms would suffer less interference from mismatches between oral and written language when learning to read and, hence, less confusion about the regularities of written English (Baratz, 1969b; Labov, 1972). Two versions of this interference hypothesis can be entertained regarding the mechanisms underlying the presumed relationship. First, the major effect of encountering such mismatches might be motivational and attitudinal changes, such that the child becomes less eager to read and less receptive to instruction, impeding academic progress. Although student attitudes were not measured in this study, we think the age differences we observed are not entirely consistent with this view. That is, we would have expected that these attitudinal changes would have been cumulative and, thus, that the magnitude of the correlations between reading scores and familiarity with SE would have grown substantially from kindergarten to second grade. There was no suggestion of that trend in our sample, although the possibility remains that a longitudinal study might reveal the expected increase over time.

Alternatively, the greater disparity between oral and written forms of English for children who are less familiar with SE might make it harder for them

to discover and learn particular correspondences between spellings and spoken words. For a child familiar only with AAVE, for instance, the oral counterpart of the written sentence "Their hands are cold" could legitimately be /dɛjr hæn a co/ ("Deir han' a' co"'). If so, this would give rise to several potentially confusing mismatches between dialects that would not be encountered by a non-AAVE-speaking child: (a) The straightforward correspondence between the phoneme /d/ and the letter *D* in SE may be harder to discover and learn because the first sound of "deir" is not spelled with the same letter used to represent the first sound of *dog* but rather is spelled with the digraph *TH*. (b) The spelling of the word *hand* contains a *D* for which there is no articulated phoneme in the child's speech because of reduction of the final consonant cluster. Likewise, (c) the *L* and the *D* in the spelling of *cold* and (d) the *R* in *are* have no corresponding sound in the child's spoken version of the word. (As teachers know well, the silent *E* in *are* is a confusing mismatch even for children who are highly familiar with SE.) (e) Because SE requires number agreement between subject and verb, the written sentence contains the verb *are* following a plural subject rather than what the AAVE-speaking child might expect to occur (i.e., the word *is* or perhaps no copula at all). (f) Because plural count nouns must be inflected in SE, the word *hands* ends with the letter *S*, marking an inflection that is not required in AAVE. To children who do not produce the affix orally, it may not be apparent why this letter is needed.

No particular mismatch, on its own, would pose a serious impediment to learning to decode, but the accumulation of such discrepancies between oral and written forms could make grapheme–phoneme correspondences seem far less regular than they are (for SE) and, hence, more difficult to master. This argument hinges, however, on the assumption that when children produce AAVE forms in their speech, their mental representations of those forms are not more similar to SE forms, especially with regard to phonological differences. That is, adult AAVE speakers who ordinarily do not produce final consonants will typically include the absent element when a vowel follows (e.g., /hol/ for the word *hold*, but /holdin/ for the word *holding*), indicating awareness of the presence of the syllable-final phoneme that is often not expressed (Labov, 1972). (In other words, they appreciate that the words *hold* and *hole* are not true homonyms in AAVE.) If this is also the case for young children, the mismatches between AAVE and written language may be more superficial than they appear at first glance because there would be a closer resemblance of the child's lexical representations of words to the printed forms of those words. To our knowledge, however, no evidence pertaining to this question is yet available.

With regard to this linguistic interference hypothesis, it should also be noted that in kindergarten and first grade, reading was correlated as strongly with grammatical as with phonological scores, regardless of whether word recognition, pseudoword decoding, or passage comprehension was the reading measure. We had anticipated that phonological dialect differences would interfere more with learning to decode and grammatical differences would interfere with other aspects of reading acquisition. In second grade, there was indeed some indication that familiarity with morphosyntactic features of SE was more closely linked to reading comprehension, as predicted, but it is puzzling that word recognition and decoding were unrelated to phonological scores. The data thus suggest that the hypothesized interference between AAVE and SE could be less specific and more pervasive than we had supposed. Nevertheless, the notion of linguistic inference remains consistent with the overall pattern of results.

Metalinguistic awareness. A final hypothesis that merits consideration is that familiarity with SE may reflect a form of metalinguistic insight that facilitates reading acquisition more generally. That is, given equal exposure to SE and equivalent oral language proficiency, some children may be more attuned to linguistic variation in their environments, whereas others are less inclined to notice or appreciate that different ways of speaking are used in some contexts and not in others and by some speakers but not others. To say that this hypothesized sensitivity, which might be termed *dialect awareness*, is metalinguistic means that it transcends the ability simply to speak and listen proficiently; rather, it involves treating language as something to be thought about, analyzed, judged, and even played with. It is well established that metalinguistic ability, especially phonological awareness but also syntactic and lexical awareness, is linked both theoretically and empirically to how readily young children learn to read (Bohannon, Warren-Leubecker, & Hepler, 1984; Committee on the Prevention of Reading Difficulties in Young Children, 1998; Tunmer & Hoover, 1992) and is itself reciprocally strengthened by the acquisition of literacy (Ehri & Wilce, 1980; Perfetti, Beck, Bell, & Hughes, 1987). Furthermore, extensive research in recent years has demonstrated that phonological awareness can be taught readily to young children (e.g., Brady, Fowler, Stone, & Winbury, 1994; Byrne & Fielding-Barnesly, 1995), and such training

has been shown to facilitate learning to read (National Reading Panel. 2000).

The number of AAVE features a child produces when imitating SE sentences, therefore, may be a marker for metalinguistic ability in young African American children, and this broad dimension of individual difference may be what underlies variation in their reading achievement (as in mainstream populations). According to this hypothesis, children with stronger metalinguistic abilities (including, but presumably not restricted to, dialect awareness) would more readily learn to read because they are more attuned to language generally rather than just because they encounter less interference between oral and written forms (although the interference and metalinguistic hypotheses are certainly not mutually exclusive). This explanation would be consistent with the equipotentiality of our phonological and grammatical scores as predictors of early reading, but without evidence regarding the relation of these scores to other metalinguistic measures, it cannot be evaluated fully.

Finally, each of the foregoing hypothesized causal mechanisms could be influenced by children's home experiences during the preschool and early school years. Parents may vary in the degree to which they expose children to AAVE and SE features in their speech at home, encourage SE usage or discourage AAVE usage in school-related activities, provide home literacy experiences that promote the learning of spelling–sound relationships, draw children's attention to language variation, and so forth. Such differences could affect children's familiarity with SE and contribute to variation in early reading achievement.

Conclusions and Implications

We have identified three possible explanations that could account satisfactorily for our finding that individual differences in familiarity with SE are strongly related to reading achievement in young, African American students: instructional variation by linguistically biased teachers, linguistic interference between oral and written dialect features, and metalinguistic influences on the development of language and reading. None of these hypotheses can be definitively confirmed or ruled out on the basis of the available evidence. Because these accounts are not mutually exclusive, it is possible that future research could provide evidence that all three mechanisms operate to produce the observed relationship between reading and dialect differences. That is, a child who is more linguistically aware, who is (therefore) more familiar with SE forms, would (therefore) encounter fewer mismatches between oral and written language and could be the recipient of adequate reading instruction from a biased teacher who regards the child positively; all of these factors could contribute to successful reading acquisition, and their reverse could contribute to lower achievement.

Of course, our findings, being correlational, are not sufficient to establish causal relations of the sort that have been hypothesized. That is, we cannot know whether familiarity with SE causally influences or merely predicts reading in this population, whether attaining metalinguistic awareness brings about (or, conversely, results from) greater knowledge of SE, whether instructional practices of linguistically biased teachers differentially affect reading achievement, and so forth. To address such questions, intervention studies would be needed to establish that some programs or activities, designed in accord with one or more of the hypotheses about how and why familiarity with SE is related to reading, are beneficial to subsequent reading achievement. The particular intervention(s) one might choose to undertake, however, would depend on one's view(s) about what is most crucial for facilitating reading acquisition by AAVE-speaking children: overcoming teacher biases, fostering children's specific knowledge about SE features, or increasing children's broader metalinguistic sensitivity to language. Only by examining the interplay of these several factors (and home experiences), we believe, can a more complete causal picture be constructed of the relationship of reading acquisition to familiarity with SE.

Although the educational implications of the findings cannot realistically be specified at this time, our demonstration of substantial associations between reading and familiarity with SE suggests that research efforts on the relationship of dialect to reading was probably abandoned prematurely 20 years ago. Examining the question anew appears to be a promising route toward reaching a stronger understanding of reading acquisition by young African American students, and ultimately for providing some new instructional approaches to enhancing achievement in this population.

References

Baratz, J. (1969a). A bidialectical task for determining language proficiency in economically disadvantaged Negro children. *Child Development, 40,* 889–901.

Baratz, J. C. (1969b). Teaching reading in an urban Negro school system. In J. C. Baratz & R. W. Shuy (Eds.),

Language differences: Do they interfere? (pp. 92–116). Newark, DE: International Reading Association.

Baugh, J. (1983). *Black street speech: Its history, structure, and survival.* Austin, TX: University of Texas Press.

Baugh, J. (1999). Considerations in preparing teachers for linguistic diversity. In C. Adger, D. Taylor, & O. Taylor (Eds.), *Making the connection: Language and academic achievement among African American students* (pp. 81–96). Washington, DC: Center for Applied Linguistics.

Bereiter, C. (1966). *Teaching disadvantaged children to read.* Englewood Cliffs, NJ: Prentice Hall.

Bohannon, J., Warren-Leubecker, A., & Hepler, N. (1984). Word awareness and early reading. *Child Development, 55,* 1541–1548.

Brady, S. A., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia, 44,* 26–59.

Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow up and a new preschool trial. *Journal of Educational Psychology, 87,* 488–503.

Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book.* New York: American Heritage.

Catts, H. W., Fey, M., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal study. *Scientific Studies of Reading, 3,* 331–361.

Cazden, C. (1988). *Classroom discourse: The language of teaching and learning.* Portsmouth, NH: Heinemann.

Committee on the Prevention of Reading Difficulties of Young Children. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Research Council, National Academy Press.

Cross, J. B., DeVaney, T., & Jones, G. (2003). Pre-service teacher attitudes toward differing dialects. *Linguistics and Education, 12,* 211–227.

Cunningham, P. M. (1976-1977). Teachers' correction responses to Black-dialect miscues which are non-meaning-changing. *Reading Research Quarterly, 12,* 637–653.

Delpit, L. (1998). What should teachers do? Ebonics and culturally responsive instruction. In T. Perry & L. Delpit (Eds.), *The real Ebonics debate* (pp. 17–26). Boston: Beacon.

Denton, K., & West, J. (2002). *Children's reading and mathematics achievement in kindergarten and first grade.* Washington, DC: National Center for Education Statistics.

Dillard, J. L. (1972). *Black English: Its history and usage in the United States.* New York: Random House.

Dreeben, R. (1987). Closing the divide. *American Educator, 11*(2), 28–35.

Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonetic structure of words. *Applied Psycholinguistics, 1,* 371–385.

Goodman, K. S., & Buck, C. (1973). Dialect barriers to reading comprehension revisited. *The Reading Teacher, 27,* 219–223.

Harber, J. R., & Beatty, J. W. (1978). *Reading and the Black English-speaking child: An annotated bibliography.* Newark, DE: International Reading Association.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Brookes.

Labov, W. (1972). *Language in the inner city: Studies in the Black English vernacular.* Philadelphia: University of Pennsylvania Press.

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: Issues in learning, teaching, and schooling* (pp. 39–68). Cresskill, NJ: Hampton.

Labov, W. (1998). Co-existent systems in African-American English. In S. Mufwene, J. Rickford, J. Baugh, & G. Bailey (Eds.), *The structure of African-American English* (pp. 110–153). London: Routledge.

Labov, W., Cohen, P., Robins, C., & Lewis, J. (1968). *The nonstandard English of Black and Puerto Rican speakers in New York City.* Final report, Cooperative Research Project 3288. Philadelphia: U.S. Regional Survey.

Labov, W., & Harris, W. (1986). De facto segregation of Black and White vernaculars. In D. Sankoff (Ed.), *Diversity and diachrony* (pp. 1–24). Philadelphia: John Benjamin.

McCardle, P., Scarborough, H. S., & Catts, H. W. (2001). Predicting, explaining, and preventing reading difficulties. *Learning Disabilities Research and Practice, 16,* 230–239.

Meir, T. (1999). The case for Ebonics as part of exemplary teacher preparation. In C. Adger, D. Taylor, & O. Taylor (Eds.), *Making the connection: Language and academic achievement among African American students* (pp. 97–114). Washington, DC: Center for Applied Linguistics.

Mufwene, S. S., Rickford, J. R., Bailey, G., & Baugh, J. (Eds.) (1998). *African American English: Structure, history, and use.* London: Routledge.

National Assessment of Educational Progress. (1995). *NAEP 1994 reading: A first look.* Washington, DC: US Government Printing Office.

National Assessment of Educational Progress. (1997). *NAEP 1996 trends in academic progress.* Washington, DC: US Government Printing Office.

National Reading Panel. (2000). *Teaching children to read: Reports of the subgroups.* Washington, DC: National Institutes of Health.

Ogbu, J. (1995). Literacy and Black Americans: Comparative perspectives. In V. Gadsden & D. Wagner (Eds.), *Literacy among African-American youth: Issues in learning, teaching, and schooling* (pp. 83–100). Cresskill, NJ: Hampton.

Owens, R. E. (1996). *Language development: An introduction* (4th ed.), Needham Heights, MA: Allyn & Bacon.

Perfetti, C., Beck, I., Bell, I., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first-grade children. *Merrill-Palmer Quarterly, 33,* 283–319.

Piestrup, A. M. (1973). *Black dialect interference and accommodation of reading instruction in first grade* (Monographs

of the Language Behavior Research Laboratory No. 4). Berkeley: University of California.

Radloff, C. F. (1991). *Sentence repetition testing for studies of community bilingualism*. Dallas, TX: Summer Institute of Linguistics and the University of Texas at Arlington.

Rickford, J. R. (1999). *African American English: Features, evolution, educational implications*. Malden, MA: Blackwell.

Rickford, J. R., & Rickford, A. E. (1995). Dialect readers revisited. *Linguistics and Education, 7*, 107–128.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.

Scarborough, H. S., Charity, A. H., & Griffin, D. (2002, June). Is unfamiliarity with "school English" (SE) related to reading achievement by African-American students? Poster session presented at the annual meeting of the Society for the Scientific Study of Reading, Chicago.

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review, 14*, 245–302.

Sims, R. (1982). Dialect and reading: Toward redefining the issues. In J. Langer & M. T. Smith-Burke (Eds.), *Reader meets author/Bridging the gap* (pp. 222–236). Newark, DE: International Reading Association.

Smitherman, G. (1977). *Talkin' and testifyin': The language of Black America*. Boston: Houghton Mifflin.

Strickland, D. (1995). Educating low-income Black children. In V. Gadsden & D. Wagner (Eds.), *Literacy among African-American youth: Issues in learning, teaching, and schooling* (pp. 69–80). Cresskill, NJ: Hampton.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*, 407–440.

Taylor, J. (1983). Influence of speech variety on teachers' evaluation of reading comprehension. *Journal of Educational Psychology, 75*, 662–667.

Tucker, R., & Lambert, W. (1969). White and Negro listeners' reactions to various American English dialects. *Social Forces, 47*, 463–468.

Tunmer, W. E., & Hoover, W. A. (1992). Cognitive and linguistic factors in learning to read. In P. B. Gough, L . C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 175–214). Hillsdale, NJ: Erlbaum.

Turner, L. D. (1949). *Africanisms in the Gullah dialect*. Ann Arbor, MI: University of Michigan Press.

Washington, J. A., & Craig, H. K. (1994). Dialectal forms during discourse of poor, urban, African-American preschoolers. *Journal of Speech and Hearing Research, 37*, 816–823.

Washington, J. A., & Craig, H. K. (1998). Socioeconomic status and gender influences on children's dialectal variation. *Journal of Speech, Language, and Hearing Research, 41*, 618–626.

Washington, J. A., & Craig, H. K. (2001). Reading performance and dialectal variation. In J. L. Harris, A. G. Kamhi, & K. E. Pollock (Eds.), *Literacy in African American communities* (pp. 147–168). Mahwah, NJ: Erlbaum.

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.

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's kindergartners: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998-99, Fall 1998*. Washington, DC: National Center for Education Statistics.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461–481.

Wolfram, W. (1974). The relationship of White southern speech to vernacular Black English. *Language, 50*, 498–527.

Wolfram, W., & Fasold, R. (1974). *The study of social dialects in American English*. Englewood Cliffs, NJ: Prentice Hall.

Woodcock, R. (1987). *Woodcock Reading Mastery Tests–Revised*. Circle Pines, MN: American Guidance Service.