■ ■ ■ Oral Language Expectations for African American Preschoolers and Kindergartners

Holly K. Craig
Julie A. Washington
University of Michigan, Ann Arbor

Fundamental to improving our clinical and empirical understanding of African American students is the development of appropriate, culture-fair assessment methods. This investigation provides descriptive information in the form of means and standard deviations for the performances of 100 typically developing African American preschool and kindergarten students on selected measures of oral language production and comprehension. The measures included four derived from free-play spontaneous language samples: a dialect density measure, mean length of C-units, frequencies of complex syntax, and number of different words. Two measures were derived from responses to nonstandardized elicitations in the form of question prompts and probes of active/passive voice construction. Performances were stable for socioeconomic status on all measures. Dialect density and number of different words varied significantly by gender. Responding to requests for information in the form of question prompts showed a statistically significant increase from preschool to kindergarten. The protocol is recommended for culture-fair assessment purposes with African American students as they begin formal education.

Key Words: language, African American, minorities, assessment

The purpose of the present investigation was to develop quantified descriptions of the oral language performances of typically developing African American children at the time of school entry. Expectations for performance variations that are considered within normal expectations for a child's age are basic to the processes involved in classroom planning and in identifying language disorders. Quantified expectations are critical to appropriate interpretation of performance variations across children so those who are typically developing are not falsely identified as language impaired nor are children with language impairments overlooked and not enrolled for needed services.

Prior studies have demonstrated that most standardized tests of language, developed for and normed on majority children, yield low performances by African American children (Hemingway, Montague, & Bradley, 1981; Washington & Craig, 1992a; Wiener, Lewnau, & Erway, 1983). The inadequacies of the Peabody Picture Vocabulary Test–Revised (PPVT-R; Dunn & Dunn, 1981) and the Test of Language Development (TOLD; Newcomer & Hammill, 1977) are notable examples of the ways in which major language tests fail to be informative for this population. On these tests, the performances of African American children cluster well below the mean (Hemingway, Montague, & Bradley, 1981; Washington & Craig, 1992a;

Wiener, Lewnau, & Erway, 1983). Further, use of tests like the TOLD-2:P (Newcomer & Hammill, 1988) can yield much higher prevalence rates of language impairment for African American children than for other segments of the U.S. population (Tomblin et al., 1997), rates that are largely uninterpretable because of concerns regarding the validity of the test scores on which these calculations are based.

As an alternative to standardized tests, assessment procedures that are less structured have been recommended for assessing the language of African American children (Seymour & Bland, 1991; Stockman, 1986, 1997; Stockman & Vaughn-Cooke, 1989). Overall, low structured tasks (especially child-centered ones for young children) such as spontaneous language sampling during free play with toys place fewer externally defined constraints and adult expectations on the child's language production (Bloom & Lahey, 1978; Brown, 1973; Lund & Duchan, 1988; Prutting, Gallagher, & Mulac, 1975). Child-centered language sampling has particular merit for assessing minority-language children in terms of reducing the impact of assumptions derived from the majority culture.

At the University of Michigan over the last few years, our research program has focused on improving understanding of the oral language use of young African American children. One priority has been to increase

knowledge of the child's use of African American English (AAE). Considerable data have been available concerning the nature of AAE and its widespread use by African Americans (Baratz, 1970; Dillard, 1972; Fasold & Wolfram, 1970; Smitherman-Donaldson, 1977; Wolfram & Fasold, 1974). However, most of this information derives from the study of adolescents and adults, so little empirical investigation has addressed dialect use by young children. This critical omission in our understanding must be addressed because many African American children at the time of school entry speak AAE to some degree (Washington & Craig, 1994; Washington, Craig, & Kushmaul, 1998), and their dialect may not be a good match to the Standard American English (SAE) of the classroom and curriculum materials.

AAE is primarily a morphosyntactic set of rule-governed variations from Standard American English (SAE), although both discourse and phonological differences can be observed as well (Cole & Taylor, 1990; Hester, 1996; Hicks, 1991; Michaels & Cazden, 1986). Washington and Craig (1994) identified 16 different morphosyntactic forms that characterized the discourse of African American children at the time of school entry. The two most frequent forms were zero copula/auxiliary (e.g., "this _ her black shoe") and the subject/verb agreement form (e.g., "now she need_ some shoes"). Whereas the amount of dialect observed in a sample of spontaneous discourse varies with the average length of communication unit (MLCU) characterizing a child's discourse (Craig, Washington, & Thompson-Porter, 1998a), it has been important to this line of inquiry to develop a measure that quantifies level of dialect use: the Dialect Density Measure (DDM). AAE forms are incorporated by children into fewer than 20% of the words of even the heaviest dialect users (Washington et al., 1998). When so much of discourse does not involve AAE forms, it is possible to pursue the non-dialectal aspects of a child's expressive and receptive language skills for potential language assessment candidates (Craig, 1996). Considered as a whole, this literature suggests that culture-neutral assessment strategies for young children would be improved by focusing on the non-dialectal components of their discourse within more child-centered contexts, such as free play with toys.

The second priority of our research program, therefore, has been to develop a language assessment protocol appropriate to young children who are speakers of AAE that includes measures derived from free-play samples and that are non-dialectal in nature. This has led to extensive examination of the performances of young African American children on non-dialectal but fairly traditional approaches to language assessment. Average oral sentence lengths are a widely used method for quantifying language stage, usually in the form of mean length of utterance (Brown, 1973; Rondal, Ghiotto, Bredart, & Bachelet, 1987; Miller & Chapman, 1981; Scarborough, Wyckoff, & Davidson, 1986; Wells, 1985) or Communication Unit (C-units; Loban, 1976) for school-age children (Scott, 1988). Craig et al. (1998a) reported mean length of C-unit (MLCU) data for African American children from low-income homes and found that values increased steadily by age and grade and that syntactic complexity predicted mean C-unit lengths at a statistically significant level.

In addition to MLCUs and frequencies of syntactic complexity as measures of expressive language, our research program has examined young African American children's responses to two comprehension tasks. Craig, Washington, and Thompson-Porter (1998b) found that responses to requests for information in the form of questions, and to probes of the distinction between active and passive sentence constructions, revealed grade effects and a positive relationship to age for young African American children from middle-income homes. Consistent with earlier proposals by Leonard and Weiss (1983) that nonstandardized elicitations can play an informative role in the evaluation of the language skills of minority children, these latter tasks have been included in the protocol. Assessing comprehension using more structured probes avoids the sampling error that may result from free play, because natural contexts may not elicit an adequate number of types or tokens of targeted constructs.

Craig and Washington (2000) examined the potential of these two expressive and two receptive language measures to distinguish children with language impairments from typically developing age-mates, all of whom were speakers of AAE. They found that the performances of the group of children with language impairments (LI) were significantly lower on each measure than chronologically age-matched (CA) controls. An additional measure derived from spontaneous language samples—number of different words (NDW) as an estimate of expressive vocabulary (Watkins, Kelly, Harbers, & Hollis, 1995)—was similarly successful in distinguishing the LI children from CA controls. These five measures, therefore, have potential for identifying African American children with language disorders from typically developing peers and warrant further study. Characterizing performance expectations for normal-language children is an important next step. Clinicians would benefit from having means and standard deviations available for each measure. This would permit a comparison of the performance of any specific African American child against those of other children who are the same age and share the same cultural-linguistic background. This investigation was undertaken to begin to contribute to this process by examining the language performances of African American children in preschool and kindergarten. Undertaking the process in the beginning grades offers a strong conceptual starting point, and the outcomes should have immediate utility for clinicians engaged in early intervention programs.

Method

Participants

The participants were 100 typically developing African American students living in Metropolitan Detroit, Michigan. African Americans constituted more than 75% of the children enrolled in the participating school district. Recruitment to our research program was initiated by the school principals' sending home a project description and consent form. The first 100 students who returned signed

consent, met criteria for typical development, and were enrolled in preschool or kindergarten were the participants for this investigation.

The preschooler mean age was 55 months ($SD = 3.6$), whereas the kindergartners were 69 months ($SD = 4.4$), and overall the children ranged in age from 47 to 78 months. The SES and gender distributions of the sample were allowed to vary. See Table 1 for the resulting distributions. Socioeconomic status was determined from one or more of the following sources: the participants' eligibility or ineligibility to participate in the federally funded free or reduced-price lunch program in their schools, their eligibility for Headstart, and/or the Hollingshead Four Factor Index of Socioeconomic Status (Hollingshead, 1975) derived from caregiver interviews. The Hollingshead Index is used to assign point scores based on the occupation, years of schooling, marital status, and gender of the child's primary caregiver(s). The point totals correspond to one of five levels designed to index a family's socioeconomic status.

Only children who appeared to be typically developing were enrolled in this investigation. The children were judged to be typically developing by their teachers and parents, and they had no history of referral to nor enrollment in special education services of any type. All of the children passed a bilateral hearing screening at 25 dB for 500, 1000, and 4000 Hz (ANSI, 1989). In addition, each child was administered the Triangles subtest of the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). Triangles is a matching task that taps a fairly general cognitive skill, is appropriate for children in this age range, and evidences no racial or cultural biases (Cole, Gay, Glick, & Sharp, 1971; Kaufman, 1973; Lampley & Rust, 1986; Palmer, Olivarez, Willson, & Fordyce, 1989; Willson, Nolan, Reynolds, & Kamphaus, 1989). Each student achieved a scaled score of 7 or more, performance within one standard deviation (−3) of the mean (10). Although the conversion of raw scores to scaled scores controlled for the effects of age and grade, it seemed important to interpretation of the present data to check that the participant pool showed no systematic

variations on this measure of cognition relative to gender and socioeconomic status. No statistically significant differences were observed for scaled scores on Triangles between males ($M = 10.0$, $SD = 2.0$) and females ($M = 10.6$, $SD = 2.4$) nor between students from low ($M = 10.4$, $SD = 2.2$) or middle ($M = 10.3$, $SD = 2.3$) socioeconomic status homes (independent samples t test for gender [2, 98]: 1.28, $p > .05$; for SES [2, 98]: 0.25, $p > .05$).

Data Collection and Analysis

Spontaneous language samples, a response to the Wh-questions task, and a task probing the active/passive voice distinction were collected from all participants. The tasks were randomly ordered in the data collection protocol to avoid systematic order effects in the data. The protocols were administered by an African American female examiner who spoke AAE to the children. The children and examiners were unacquainted, consistent with most assessment contexts.

Collection and Analysis of the Spontaneous Language Samples. Fifteen- to 20-minute spontaneous language samples were collected during dyadic free play using action figures, dolls, and the Fisher-Price school. The children selected one toy set for the free-play interaction in an attempt to control for potential interest levels across children. Each child wore an individual microphone, and the samples were audio recorded using a microphone mixer.

The free-play language samples were transcribed orthographically using the segmentation criteria of Loban (1976) for C-units, defined as independent clauses plus their modifiers in the form of coordinate, subordinate, and embedded clauses. This required segmenting successive main clauses linked by simple coordinate conjunctions (*and*, *but*, *or*) into separate C-units if the second clause included a subject. When the second clause elliptically omitted the subject, the two clauses were considered a single C-unit. For example, "he's fitna get up and shoot 'em all" was considered a single C-unit, whereas the turn "it supposed to stay in so when they come in it go pow!/ and it chop them" was segmented into two C-units as indicated by the /. Loban's criteria included nonclausal utterances in the C-unit corpora if they were responses to prior adult questions. Similarly, we included single-word stereotypical acknowledgements to prior adult comments and child productions of "what?" functioning as a contingent query. Other potential single-word forms, particularly "wow!" or other fragments functioning as exclamations, were not included in the C-unit corpora.

C-units were transcribed into CHAT files consistent with the conventions of the Children's Data Exchange System (CHILDES; MacWhinney, 1994). The first 50 wholly intelligible C-units were identified and represented a standard corpus for the four expressive language measures. All children produced 50 intelligible C-units in the 15- to 20-minute sampling periods.

The samples were scored for the amounts of dialect, MLCU, amounts of complex syntax, and number of different words. The CLAN programs of CHILDES automatically generated the frequencies of AAE and

TABLE 1. Numbers of subjects by grade and socioeconomic status (SES), and by grade and gender.

	Preschool	Kindergarten	Total
SES			
Low	52	4	56
Middle	16	28	44
	68	32	100
Gender			
Male	34	11	45
Female	34	21	55
	68	32	100

Note. The four kindergartners from low-income homes were all females.

complex syntax codes using the frequency command (FREQ) and the average C-unit length in words from the mean length of turn command (MLT). Each measure is discussed in Craig and Washington (2000) and is described briefly below.

Dialect Density. Each free-play sample was scored for AAE using Washington and Craig's definitions (1994) for 16 features produced by preschoolers. The linguistic contexts in which some of these features might be produced are readily apparent. For example, inclusion or exclusion of the subject/verb agreement marker ("she keep getting stuck") is relatively easy to detect; this is a linguistic context specific to the subject/verb agreement feature, regardless of whether the form is produced or not. The contexts for some features, however, are only apparent once the feature has been produced—for example, invariant *be* ("I be pulling it"). This is a nonspecific linguistic context in which other verb forms could have been produced, depending upon the intent of the child. It would be useful to quantify amounts of dialect produced by children, but this requires the determination of an appropriate base to be used in the calculation. Whereas the features vary in the transparency of their surrounding linguistic contexts, opportunities cannot be defined at the level of the features themselves because a potential confound exists across the set of features between the type of AAE and its likelihood of being detected. Alternatively, a Dialect Density Measure (DDM) was calculated for each participant by dividing the frequencies (tokens) of AAE in their first 50 intelligible C-units by the number of words (tokens) in the same samples (Craig et al., 1998a; Craig & Washington, 2000). The DDM was developed to control for the positive relationship between increased C-unit lengths and opportunities for AAE, and it is not dependent upon the observer's ability to ascribe intent to specific linguistic contexts.

Average C-Unit Lengths. The mean length in words of the first 50 wholly intelligible C-units (MLCU) was calculated for each participant.

Syntactic Complexity. Each C-unit was scored for the presence of complex syntax (Csyn), using Craig and Washington's (1994) taxonomy. Types of complex syntax ranged from simple infinitival relationships (e.g., "it used to have a telephone" coded as an infinitive with same subject) to more complex unmarked infinitives (e.g., "you made him fall"), clauses joined by a variety of more cognitively advanced subordinate conjunctions (e.g., "when it rain we put it in the house"), and relative clauses (e.g., "actually that's the one I was talkin' about"). More than one complex syntax code was possible per C-unit. Connectives in C-unit initial positions that served as pragmatic connectives to link discourse between speakers, and those occurring in initial positions that were turn internal, were not scored as conjunctions for the purposes of this complex syntax analysis. An example of a pragmatic connective: ADU: "you pull the rope back in." CHI: "*and* I think this door can open." An example of a turn internal connective that linked consecutive C-units but was not scored for complex syntax: CHI: "you be the teacher/ *and* I'll be the kids."

Expressive Vocabulary. The number of different words (NDW) in the 50 C-unit samples was calculated as a measure of lexical diversity and expressive vocabulary. The FREQ command of CLAN automatically generated word lists. These lists were then edited so that morphological variations in the form of number and tense markers on regular nouns and verbs were ignored. Irregular forms of nouns and verbs—for example, *leaf*, *leaves* and *sit*, *sat* were treated as separate lexical forms. However, regular forms of nouns and verbs—for example, *girl*, *girls* and *walks*, *walked*, and *walking*—were treated as the same noun or verb lexical root and not scored as different words although the FREQ command displays them as different types. These bound morphemes may be variably included or excluded when a child is speaking AAE, so it seemed appropriate to represent this aspect of their production within the dialect analyses rather than this lexical analysis.

Comprehension of Requests for Information. This task involved presenting two activity pictures to the children—depictions of barbecuing and snow shoveling—and probing for responses to questions.[1] For each picture the examiner asked 12 questions using AAE: What this (is)? What he doin'? Who (object) this? Who this? How many (objects) in the picture? Where this? How long will it take to (perform action)? Why he (perform action)? How far he (perform action)? How he (perform action)? How often he (perform action)? and When this happenin'? The order of presentation of each prompt for each picture was randomized for each participant, although pilot work indicated that order effects were not apparent for nonrandomized trials (see Craig et al., 1998b). Scoring assigns full credit (3 points) if the child produces the target response; 2 points if the child responds to the pragmatic intent of the specific request for information but uses a nonspecific referent or misnames the referent (e.g., ADU: "how he movin' the snow?" CHI: "with a lawnmower"); 1 point if the child responds but to a potentially different question (e.g., ADU: "how often they barbeque?" CHI: "because they hungry"); and 0 points if the child says something unrelated, "I don't know," or does not respond. The total possible score was 72 points. (See Craig et al., 1998b.) A computerized scoring program generated total scores and error analyses automatically.

Comprehension of Active/Passive Sentences. This task explored the children's word order strategies for comprehension of reversible active and passive sentence constructions (RevS) using a forced-choice picture-pointing task. The stimulus set of picturable agents and actions was pretested with each participant for familiarity, and all participants readily identified the referents. RevS consists of 30 spoken prompts, three for each of 10 pairs of pictures

in the form of black line drawings on 5 × 7 cards. The prompts use common nouns and verbs that were selected for their likelihood to occur bi-directionally in real life experiences. An example was "the mom hug the baby" (target active voice trial); "the baby hug the mom" (foil active voice trial); "the baby was hugged by the mom" (passive trial). Earlier pilot work revealed a statistically significant advantage when active trials consistently preceded the passive ones (Craig et al., 1998b). Both the order of the picture pairs and the order of the spoken prompts, therefore, were randomly determined for each participant. The total point score was 20, and one point was assigned for a match between the target active voice trial and one point for a match between the passive voice trial. A match between the prompt and the child's picture selection on the passive voice trial, however, was credited only when the child correctly matched both active voice trials as well (for additional discussion see Craig et al., 1998b). A computerized scoring program generated total scores and error analyses.

Reliability

Reliabilities were established for the spontaneous language samples by randomly selecting one of the 50 C-units in each transcript as a starting point and then having an independent observer retranscribe approximately 15% of the data as a whole. For transcription, a point-to-point comparison at the level of the morpheme was high (91%) when the number of morphemes in agreement was divided by the number of agreements plus disagreements. C-unit segmentation for these same samples was also high (96%).

Fifteen transcripts, representing 15% of the corpus, were randomly selected and re-coded by an independent observer. Point-to-point comparisons were calculated for each scoring system by dividing the number of agreements by the number of agreements plus disagreements. The percentages of agreement for AAE types was 98% and for AAE tokens was 94%. The percentages of agreement for types of complex syntax was 95% and for tokens was 89%. These analyses indicated high levels of agreement across transcription and scoring portions of the data reduction.

Results

Each of the measures was examined for systematic variations using multivariate general linear models that combined both analysis of variance (ANOVA) and regression models. Effect sizes (ES, Eta Squared) were small for all analyses. See Table 2. Despite the small ES for all variables, gender was found to be statistically significant (.21, $p = .002$).

All but five of the children used AAE during the 50 C-unit samples. The children who did not use AAE within the first 50 C-units did so later in their free-play samples ($n = 4$) or while describing pictures ($n = 1$). Overall, every child produced one or more of the AAE forms, and 95% of the participants demonstrated use of AAE within their first 50 C-units.

There were no significant interaction effects for the

TABLE 2. Multivariate analyses of variance _F_ ratios and effect sizes for socioeconomic status (SES), grade, and gender.

Effect	Wilks' Lambda	_F_	_p_	Eta Squared
SES	.94	.81	.57	.05
Grade	.94	1.03	.41	.07
Gender	.79	3.85	.002	.21
SES Grade	.96	.68	.67	.04
SES Gender	.98	.35	.91	.02
Grade Gender	.96	.61	.71	.04

Note. An SES/Grade/Gender analysis was not possible because there were no male kindergartners in the sample. The _df_ were 6, 88 for all analyses.

DDM relative to grade, gender, and SES, nor main effects for grade [$F(1, 93) = 0.16$, $p > .05$] or SES [$F(1, 93) = 1.87$, $p > .05$]. However, there were significant main effects for gender [$F(1, 93) = 13.58$, $p = .000$]. Boys produced significantly more dialect tokens (Mean DDM = .061) than girls (Mean DDM = .036). Table 3 presents means and standard deviations for DDM by gender. This ES was small for DDM but statistically significant (.13, $p = .000$).

Two forms were the most frequently used types of AAE both within and across children: the zero copula/auxiliary ("this _ where you live") and the subject-verb agreement ("that's how it go__") forms. There were no statistically significant differences by gender in the percentage frequencies of occurrence of either the zero copula/auxiliary feature [$t(86) = .06$, $p > .05$] or the subject/verb agreement feature [$t(73) = 1.57$, $p > .05$], indicating that despite gender differences in the amounts of dialect produced, the types of AAE the boys and girls used were the same. Figure 1 presents percentage frequencies of each type of AAE distributed across the 100 students (% of participants) as well as the percentage frequencies of each type relative to the amount of dialect produced by each child (% of AAE). The former indicates the degree to which a particular feature is represented in the discourse of preschool children and kindergartners, whereas the latter reflects the extent to which each feature is used in conversation.

The double copulas/auxiliaries/modals form ("I'm is the last one") was not widely dispersed across the sample of participants, but for those children who used this type of AAE it was relatively frequent in their discourse. The other types of AAE were used by fewer than one quarter of the students. Further, within the children's samples, these other types of AAE represented fewer than one quarter of the instances of AAE production. See Figure 1.

Mean C-unit lengths and mean frequencies of complex syntax evidenced no statistically significant relationships. For MLCU, there were no interaction effects and no statistically significant main effects for grade [$F(1, 93) = 0.28$, $p > .05$], gender [$F(1, 93) = 1.81$, $p > .05$], or SES [$F(1, 93) = 0.02$, $p > .05$]. Most students, 93%, produced

TABLE 3. Means (*M*) and standard deviations (*SD*) for the dialect density measure (DDM) by gender, mean length of C-units (MLCU), frequencies of complex syntax (Csyn), number of different words (NDW) by gender, responses to Wh-questions task (Wh-q) by grade, and the reversible sentences task (RevS).

		M	*SD*	−1 *SD*	−1.25 *SD*
DDM	Male	.061*	.038	na	na
	Female	.036*	.022	na	na
MLCU		3.15	.72	2.43	2.25
Csyn		4.3	3.2	1.1	0.3
NDW	Male	70.6**	16.5	54.1	50.0
	Female	79.5**	14.8	64.7	61.0
Wh-q	Preschool	53.2***	5.3	47.9	46.6
	Kindergarten	58.0***	5.3	52.7	51.4
RevS		12.3	3.8	8.5	7.5

Note. na = not appropriate; not a clinical measure.
* *p* = .000, ** *p* = .017, *** *p* = .036

FIGURE 1. Percentage frequencies of AAE types across the 100 participant sample (% of participants) and within the 50 C-unit corpora of each participant (% of AAE). COP = zero copula or auxiliary. SVA = subject-verb agreement. NEG = multiple negation. POS = zero possessive. UPC = undifferentiated pronoun case. PST = zero past tense. FSB = fitna/sposeta/bouta. PRO = appositive pronoun. IBE = invariant *be.* ZPL = zero plural. ZTO = zero *to.* ART = indefinite article. ING = zero *–ing.* AIN = ain't. MOD = double copula/auxiliaries/modals.
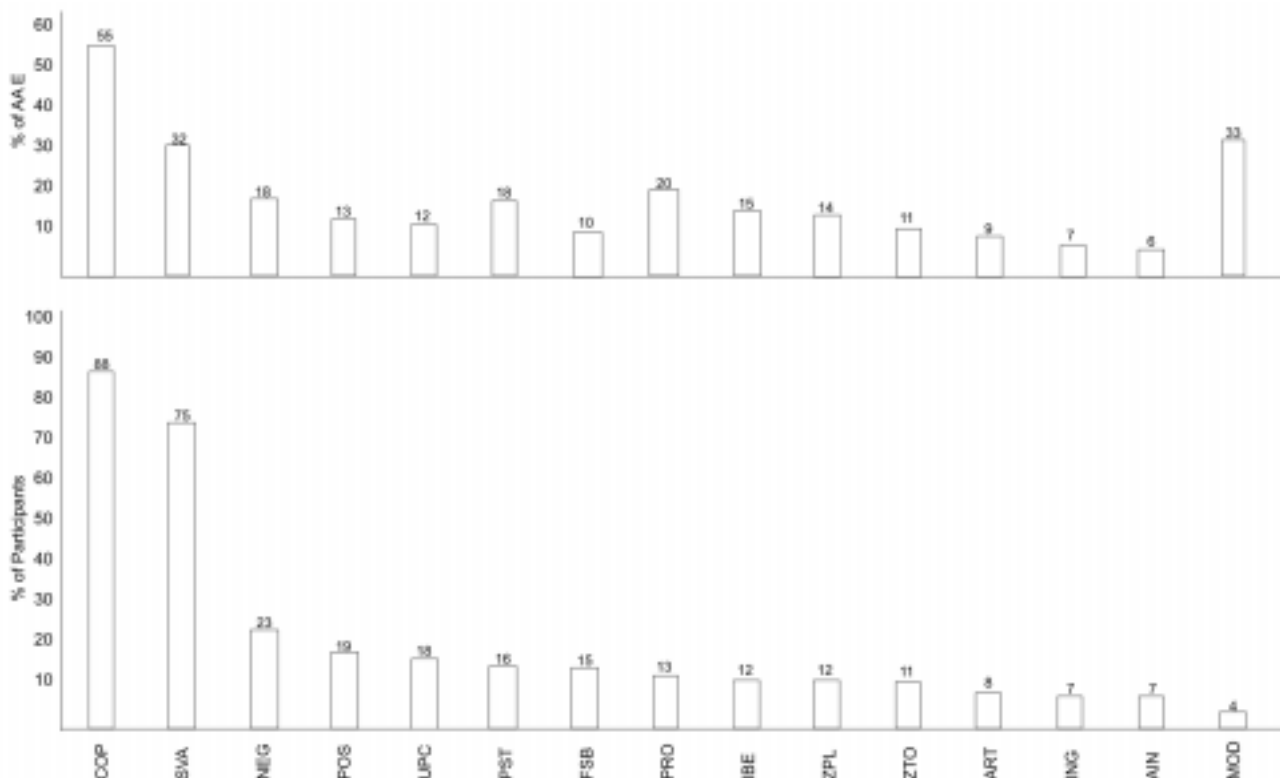
one or more instances of complex syntax within their 50 C-unit free-play samples. For Csyn, there also were no interaction effects nor main effects for grade [$F(1, 93) = 0.04$, $p > .05$], gender [$F(1, 93) = 1.81$, $p > .05$], or SES [$F(1, 93) = 0.65$, $p > .05$]. Table 3 presents the means and standard deviations for these two production measures. The values are presented as totals in the absence of significant main effects relative to grade, gender, or SES. In addition, Figure 2 presents the percentage frequencies of occurrence of each type of complex syntax. Considered across the cohort of 100 participants (% participants), the samples of approximately half of the students evidenced conjunctions, either in the form of a non-infinitive wh-conjunction (e.g., "I forgot how old she is") or a coordinate (e.g., "Robin be doing this and kicking") or subordinate conjunction (e.g., "he use it so nobody can get him"). Simple infinitives referencing the same subject (ISS; e.g., "they don't have no room to sleep in") also were produced by almost half of the students. About one-third of the students used noun phrase complements (e.g., "yeah I know him is right there") in their 50 C-unit free-play samples, but other types of complex syntax were used by one quarter or fewer of the participants. In contrast, considered relative to the 50 C-unit base for each child (% Csyn), the distribution within the children's free-play samples was relatively flat.

The number of different words evidenced no significant interaction effects and no significant main effects for grade [$F(1, 93) = 0.00$, $p > .05$] or SES [$F(1, 93) = 0.27$, $p > .05$]. However, there was a significant main effect for gender on the NDW measure [$F(1, 93) = 5.96$, $p = .017$]. Females produced a statistically significant greater number of different words (Mean = 79.5) than males (Mean = 70.6) in their 50 C-unit samples. See Table 3. The ES was again small for NDW but statistically significant (.06, $p = .017$).

Performances on the Responses to Wh-questions task (Wh-q) revealed no significant interaction effects and no significant main effects for gender [$F(1, 93) = 0.22$, $p > .05$] or SES [$F(1, 93) = 2.04$, $p > .05$]. However, there was a statistically significant main effect for grade [$F(1, 93) = 4.52$, $p = .036$]. Kindergartners achieved significantly higher scores (Mean = 58.0) than preschoolers (Mean = 53.2). Table 3 presents the mean values for Wh-q by grade. The ES for Wh-q by grade was small but statistically significant (.05, $p = .04$).

Reversible sentences evidenced no significant interaction effects and no significant main effects [$F\text{grade}(1, 93) = 1.12$, $p > .05$; $F\text{gender}(1, 93) = 0.90$, $p > .05$; $F\text{SES}(1, 93) = 1.33$, $p > .05$]. See Table 2. Most students' scores were the result of correctly responding to the active trials. Only approximately 33% of the participants identified the passive prompts on more than 5 of the 10 trials, a level considered above chance.

FIGURE 2. Percentage frequencies of Csyn types across the 100 participant sample (% of participants) and within the 50 C-unit corpora of each participant (% of SYN). CON = Coordinate and subordinate conjunctions. NIW = simple noninfinitive Wh-c lause. ISS = infinitive same subject. NPC = noun phrase complement. GER = gerunds and participles. REL = relative clause. LET = let(s)/lemme and infinitive. UNI = unmarked infinitive. IDS = infinitive with a different subject. WHI = Wh-infinitive clause. Tag = tag questions.
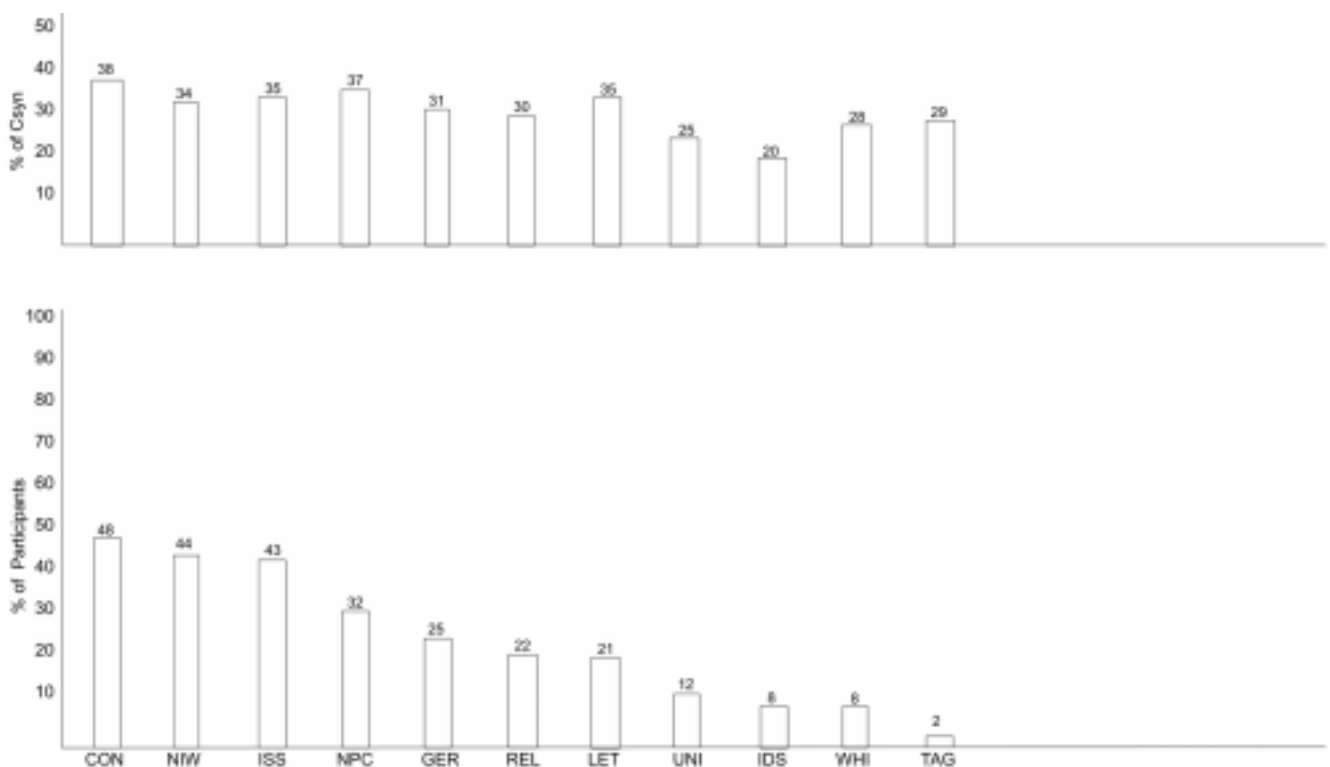
Discussion

This study reports means and standard deviations for typically developing African American preschoolers and kindergartners on oral language tasks that represent fairly traditional approaches to describing language skills. The investigation extends earlier work that showed that measures like these increase systematically with age and grade (Craig et al., 1998a, 1998b). The present study clarifies the relationships among grade, gender, and SES for each of these measures. Effect sizes for all variables in these multivariate analyses were small. However, despite small effect sizes, significant relationships were detected on the DDM and NDWs for gender and on the Wh-q for grade. The finding of significance despite small effects indicates that these analyses had enough power to identify significant but small effects when present. Further, the participant sample included 100 cases, the nonsignificant relationships evidenced negligible effect sizes, and the amount of variance associated with the outcome measures was relatively small for all measures except Csyn (see *SD* in Table 3). Although future research will be needed to confirm the relationships observed in this study, considered together, the current data indicate that these findings are valid and can be used to begin to create quantified expectations for oral language performances of young African American children.

Recently, Craig and Washington (2000) demonstrated how the non-dialectal portion of this protocol could be used to identify African American children with language impairments. Using a standard of below chronological age-mate performances on two or more of the MLCU, Csyn, NDW, Wh-q, and RevS measures, all of the 24 children with language impairments were identified as language impaired in this earlier work. Therefore, the values reported in the current investigation provide potentially useful information to clinicians and researchers for identifying or confirming the presence of language disorders in young African American children. It will be a fairly straightforward task to compare performances of individual students on these tasks against the –1 or –1.25 standard deviations below the mean recommended in the clinical literature (Aram, Morris, & Hall, 1993; Tomblin, Records, & Zhang, 1996). The Triangles subtest of the K-ABC (Kaufman & Kaufman, 1983), the Peabody Picture Vocabulary Test-III (Dunn & Dunn, 1997), and the Arizona Articulation Proficiency Scale (2nd ed.; Fudala & Reynolds, 1986) offer nondiscriminatory assessments of generalized cognition (Cole et al., 1971; Kaufman, 1973; Lampley & Rust, 1986; Palmer et al., 1989; Willson et al., 1989), receptive single-word vocabulary (Washington & Craig, 1999), and articulation skill (Washington & Craig, 1992b) for young African American children. Considered together with the outcomes of the current investigation, the clinician now has an appropriate and effective repertoire of language assessment measures for African American students when they enter school. Whereas these kinds of measures have a long history in our profession, they should be useful immediately to practitioners skilled in the collection and analysis of language samples and in the elicitation of non-standardized probes. They require no special training beyond a sound knowledge of language sampling and analysis and, thus, recommend themselves in this regard.

As part of our research program, we have been able to use the information from this investigation to assist classroom teachers with classroom planning and to identify children appropriate for referral to special education services. As background, Craig and Washington (2000) demonstrated that this protocol has excellent sensitivity (1.00) and specificity (.86) for African American children with language impairments compared to chronological age and MLCU-matched peers. The criteria we used in that prior investigation required low scores on two or more of the five measures. That investigation employed a matched participant research design, so there was no independent cohort of typically developing children at each grade from which to derive means and standard deviations by measure.

The current study provides these needed means and standard deviations on a sample of 100 African American preschoolers and kindergartners. With this information, we have been able to assist clinicians and teachers, often unsure of the status of an African American child's language when the child is a speaker of AAE, in determining which children approximate normal developmental expectations for oral language and which children warrant continued observation. For example, two preschoolers we will refer to as "Jamal" and "Gerard" who were not a part of the present investigation were referred to us. Both boys were from low-SES homes, lived in the same community, and were enrolled in the same school district. Both boys were heavy dialect users (DDM: Jamal = .118, DDM: Gerard = .126) in that their instances of AAE production relative to the number of words they spoke were considerably higher than the mean DDM for boys in this investigation (see Table 3). Jamal's oral language skills approximated the mean for the sample in the current investigation, in that the values obtained for him were within one standard deviation on all measures: MLCU = 3.22, Csyn = 7, NDW = 69, Wh-q = 54, and RevS = 10. In contrast, Gerard's oral language skills appeared weak, in that the values obtained for Gerard were more than one standard deviation below the mean (see Table 3) for MLCU and Csyn (MLCU = 2.38 and Csyn = 1, respectively) and more than 1.25 standard deviations below the mean for Wh-q and RevS (Wh-q = 30 and RevS = 1, respectively). Only Gerard's performance on the NDW measure (NDW = 66) approximated the mean we obtained for our typically developing cohort in the current study. These profiles indicate that Jamal will be able to participate in oral language classroom tasks like his peers. Gerard, however, will have more difficulty expressing himself than his peers and even more difficulty with oral language comprehension. Teachers will be able to use this type of information to provide supportive classroom instruction for students like Gerard, and indeed Gerard required special education support services.

As a whole, the protocol yielded quite stable performance outcomes, with few statistically significant variations by grade or in terms of potentially important social status characteristics of the children. Whereas the present

investigation was limited to one geographic area and to an urban, predominantly middle-class community, it will be important for future research to determine how broadly these outcomes can be generalized to other parts of the country. Further, other measures may be equally as informative or indeed better than those selected in the current study, but this awaits further research.

The failure to find SES differences is surprising. Systematic variations relative to SES were not apparent even for dialect, although in some of our earlier work (Washington & Craig, 1998) we found a significantly greater frequency of AAE tokens for children from low- than from middle-income families. In the earlier work, we examined opportunities for occurrence of AAE by correlating AAE tokens and MLCU in words, finding a nonsignificant relationship. In Craig et al. (1998b) we controlled more specifically for the potential confound between numbers of words and AAE opportunities by creating a density measure: frequency of AAE tokens/words in sample. The current study is the first opportunity in our research program to re-examine dialect density, defined as it was by Craig et al. (1998b), relative to SES. It may be that SES differences are observable on a surface level, but not when co-variables are controlled.

Alternatively, the advantages represented by the greater resources associated with middle- as opposed to low-income status reportedly affect a host of child development measures (Hart & Risley, 1995; McLoyd, 1998). In the current study, despite family social status differences, all of the children resided in the same community and were enrolled in the same school district. Like the students in this investigation, the participants in the earlier examination of dialect (Washington & Craig, 1998) resided in Metropolitan Detroit. However, in the earlier study, the poor students were enrolled in a different school district within the Detroit Metropolitan area, and the demographics of that community were overwhelmingly poor. The middle-income families were also from Metro Detroit but resided in a predominantly middle-income community. The low- and middle-income students in the current investigation were from the latter community, which was largely middle-income. Perhaps the socioeconomic disadvantages for the children of low-income families are mitigated when the community in which the family resides and the schools attended have the resources available associated with middle-income status that characterize the majority of the community. For example, in other studies (see McLoyd, 1998), family poverty and community poverty each acted independently to predict lower levels of school achievement, controlling for various parent and family characteristics. The merits of this explanation are not testable within this investigation but warrant further study given that so many African American children are growing up in low-income families and so many African American students do not fare well in school.

The Wh-questions task was the only measure yielding a significant increase in performances by grade. It is not surprising that these performance outcomes reflected so few grade effects overall. Quantitative measures like these are fairly general assessments of oral language, and

although they distinguish outcomes associated with language impairments, they may not be highly sensitive to normal variations (Klee & Fitzgerald, 1985; Watkins et al., 1995). Further, as observed by others, once typically developing children are beyond the earliest stages of oral language acquisition, the spans needed to detect substantive changes in language become longer and often are not apparent between consecutive chronological years of age or grade (Nippold, 1988; Scott, 1988). The current findings are consistent with these prior observations.

It is not clear from the current data why the Responses to Wh-questions task was an exception to the general trend in that it did show a statistically significant increase in scores between the preschoolers and kindergartners. It was beyond the scope of the current investigation to collect detailed case history information on each child, so it is not known whether the kindergartners had prior preschool experience. If they did, the significantly greater scores for the kindergartners might reflect the positive effects of early school experience on learning to respond to requests for information. It is the case that early childhood programs can have a positive and lasting effect on reading and mathematics achievement, grade retention, and assignment to special education (Campbell & Ramey, 1995). Others have observed that requiring responses by children to questions of adults is not characteristic of African American homes (Anderson-Yockel & Haynes, 1994). Accordingly, the preschoolers in our study may have had limited experience with this discourse routine prior to school entry. Alternatively, this task taps extant world knowledge more than the other tasks in the protocol, and the increased scores of the kindergartners may be attributable simply to the addition of another year of life experiences compared to the preschoolers.

In addition to a systematic relationship between grade and responding to requests for information, significant gender differences were observed for the dialect density measure and number of different words. The finding of greater AAE production by boys is consistent with prior research for both children (Washington & Craig, 1998) and adults (Chambers, 1992; Wolfram, 1969). Boys produced AAE at approximately twice the level of the girls, a finding that may relate to differences in socialization practices for boys and girls (see Washington & Craig, 1998). Although both boys and girls evidence heaviest use of zero copula and auxiliaries as well as variable inclusion and exclusions of the subject-verb agreement marker, the higher DDMs for boys indicate that their discourse will reflect relatively high levels of these two forms of AAE. These two forms of AAE have been proposed as part of a tense-based clinical marker for children with specific language impairment (Leonard, 1995; Loeb & Leonard, 1991; Rice & Wexler, 1996; Rice, Wexler, & Cleave, 1995). Of theoretical import, the current data indicate that the zero copula/auxiliary and subject/verb agreement types of AAE are a critical test of the crosslinguistic generalizability of tense as a clinical marker and should be pursued. Of practical import, the current data indicate that to the extent that zero copula/ auxiliary and subject/verb agreement types of AAE influence our identification of children with language disorders,

African American boys are at particular risk for misidentification. At a surface level, the two most frequent types of AAE, the zero copula/auxiliary and subject/verb agreement features, are indistinguishable from deletions of tense markers in the discourse of children with language disorders. Considerable additional research is needed to specify the principles that distinguish dialect from disorder. This currently unresolved problem underscores the utility of using non-dialectal measures like MLCU, Csyn, NDW, Wh-q, and RevS in conjunction with the K-ABC, PPVT-III, and APPS for language assessment purposes with preschool and kindergarten African American students.

Girls produced significantly higher mean number of different words than boys. Gender differences were not observed for the other non-dialectal aspects of language production—including amounts of complex syntax, average C-unit lengths, responding to Wh-question prompts—nor to reversible sentences. It is beyond the scope of the present study to determine why NDWs evidenced a gender effect. Perhaps differences between boys and girls in lexical diversity during conversation are an early form of gender-based differences in conversational styles reported for adults. For example, Nordenstam (1992) reported greater word production rates for female than for male adults, and NDWs may be a related measure. Unfortunately, very little information about gender-based differences in conversational styles is available for young children, and the current findings await interpretation until this larger context is available.

It is interesting that every child in this investigation spoke AAE. However, it is not clear how this heritage language affects school success. It is the case that African American students are at risk for academic failure. They perform lower than majority peers on standardized tests of academic achievement, are over-represented on the nation's special education caseloads (Statistical Profile of Special Education in the United States, 1994; Nettles & Perna, 1997), and perform significantly lower than non-minority students in vocabulary, reading, writing, science, math, and geography (U.S. Department of Education, National Center for Education Statistics, 1998). This longstanding "Black-White Test Score Gap" (Jencks & Phillips, 1998) is apparent at the time of school entry, especially in language and literacy skills (Nettles & Perna, 1997). The gap widens as early as the first quarter of first grade (Entwisle & Alexander, 1988) and increases through twelfth grade (Phillips, Crouse, & Ralph, 1998). Further, 6 of the 10 indicators of success or failure in reading identified by the National Research Council report on preventing reading difficulties in children relate to language skills (Snow, Burns, & Griffin, 1998). The first step in improving academic outcomes for African American students should be to improve our knowledge of the language skills that African American children bring to classrooms. This investigation begins to contribute to this research imperative.

The data indicate that the typically developing African American child will enter formal schooling with the following oral language characteristics:

1. The student will be a speaker of AAE.

2. The student's AAE will be characterized by copula and auxiliary deletions and inconsistent omissions of number agreement markers on verbs. These effects will be more pronounced for boys.

3. On average, the student will speak in C-units ranging from approximately 2.5 words to 4 words in length.

4. The student will use conjunctions and simple infinitives to combine clauses.

5. If the student is a boy, he will likely use a less diverse expressive vocabulary than girls during spontaneous conversation.

6. The student will be responsive to requests for information.

7. The student will likely understand active statements but not passive ones.

Teachers and speech-language pathologists should find the above information useful as they help African American students succeed in classroom contexts. The information provided in this study should contribute to the knowledge base available to practitioners but underscores that considerable additional research is needed to improve our understanding of the oral language skills of African American students.

Acknowledgments

References

American National Standards Institute. (1989). *Specifications for audiometers* (ANSI S3.6-1989). New York: ANSI.

Anderson-Yockel, J., & Haynes, W. O. (1994). Joint book-reading strategies in working-class African American and white mother toddler dyads. *Journal of Speech and Hearing Research, 32*, 583–593.

Aram, D., Morris, R., & Hall, N. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research*, *36*, 580–591.

Baratz, J. C. (1970). Educational considerations for teaching standard English to Negro children. In R. W. Fasold & R. W. Shuy (Eds.), *Teaching Standard English in the inner-city* (pp. 20–39). Washington, DC: Center for Applied Linguistics.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: John Wiley and Sons.

Bracken, B. A. (1986). *Bracken concept development program.* San Antonio, TX: Psychological Corporation.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high-risk African American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal, 32,* 743–772.

Chambers, J. K. (1992). Linguistic correlates of gender and sex. *English World-Wide, 13*(2), 173–218.

Cole, M., Gay, J., Glick, J., & Sharp, D. W. (1971). *The cultural context of learning and thinking.* New York: Basic Books.

Cole, P. A., & Taylor, O. L. (1990). Performance of working-class African American children on three tests of articulation. *Language, Speech, and Hearing Services in Schools, 21,* 171–176.

The Council for Exceptional Children. (1994). *Statistical profile of special education in the United States.* Reston, VA: Author.

Craig, H. (1996). The challenges of conducting language research with African American children. In A. Kamhi, K. Pollock, & J. Harris (Eds.), *Communication development and disorders in African American children.* Baltimore: Paul Brooks.

Craig, H., & Washington, J. (1994). The complex syntax skills of poor, urban, African American preschoolers at school entry. *Language, Speech, and Hearing Services in Schools, 25,* 181–190.

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

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

Craig, H., Washington, J., & Thompson-Porter, C. (1998b). Performances of young African American children on two comprehension tasks. *Journal of Speech, Language, and Hearing Research, 41,* 445–457.

Dillard, J. (1972). *Black English.* New York: Random House.

Dunn, L., & Dunn, L. (1997). *The Peabody Picture Vocabulary Test–III (PPVT-III).* Circle Pines, MN: American Guidance Service.

Dunn, L., & Dunn, L. (1981). *The Peabody Picture Vocabulary Test–Revised.* Circle Pines, MN: American Guidance Service.

Fasold, R., & Wolfram, W. (1970). Some linguistic features of Negro dialect. In R. W. Fasold & R. W. Shuy (Eds.), *Teaching Standard English in the inner-city* (pp. 41–86). Washington, DC: Center for Applied Linguistics.

Fudala, J. B., & Reynolds, W. M (1986). *Arizona Articulation Proficiency Scale* (2nd ed.) Los Angeles: Western Psychological Services.

Entwistle, D. R., & Alexander, K. L. (1988). Factors affecting achievement test scores and marks of black and white first graders. *The Elementary School Journal, 88,* 449–472.

Hart, B., Risley, T. (1995). *Meaningful differences.* Baltimore, MD: Paul H. Brookes.

Hemingway, B. L., Montague, J. C., & Bradley, R. H. (1981). Preliminary data on revision of a sentence repetition test for language screening with Black first-grade children. *Language, Speech, and Hearing Services in Schools, 12,* 145–152.

Hester, E. (1996). Narratives of young African American children. In A. Kamhi, K. Pollock, & J. Harris (Eds.), *Communication development and disorders in African American children* (pp. 227–245). Baltimore: Paul H. Brookes.

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

Hollingshead, A.B. (1975). *Four factor index of social status.* (Available from author, P.O. Box 1965 Yale Station, New Haven, CT).

Jencks, C., & Phillips, M. (1998). *The Black-White test score gap.* Washington, DC: Brookings Institute.

Kaufman, A. (1973). Comparison of the performance of matched groups of black children and white children on the Wechsler Preschool and Primary Scale of Intelligence. *Journal of Consulting & Clinical Psychology, 41*(2), 186–191.

Kaufman, A., & Kaufman, N. (1983). *Kaufman Assessment Battery for Children: Interpretive manual.* Circle Pines, MN: American Guidance Service.

Klee, T., & Fitzgerals, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language, 12,* 251–269.

Lampley, D. A., & Rust, J. O. (1986). Validation of the Kaufman Assessment Battery for children with a sample of preschool children. *Psychology in the Schools, 23,* 131–137.

Leonard, L. B. (1995). Functional categories in the grammars of children with specific language impairment. *Journal of Speech and Hearing Research, 38,* 1270–1283.

Leonard, L., & Weiss, A. (1983). Application of non-standardized assessment procedures to diverse linguistic populations. *Topics in Language Disorders, 3*(3), 35–45.

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

Loeb, D. F., & Leonard, L. B. (1991). Subject case marking and verb morphology in normally developing and specifically language-impaired children. *Journal of Speech and Hearing Research, 34,* 340–346.

Lund, N. J., & Duchan, J. F. (1988). *Assessing children's language in naturalistic contexts.* Englewood Cliffs, NJ: Prentice-Hall.

MacWhinney, B. (1994). *The CHILDES project: Tools for analyzing talk* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

McLoyd, V. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53,* 185–204.

Michaels, S., & Cazden, C. (1986). Teacher/child collaboration as oral preparation for literacy. In B. Schieffelin & P. Gilmore (Eds.), *The acquisition of literacy: Ethnographic perspectives* (pp. 132–154). Norwood, NJ: Ablex.

Miller, J., & Chapman, R. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research, 24,* 154–161.

Nettles, M. T., & Perna, L.W. (1997). *The African American education data book* (Vol. II: Preschool through high school). Ann Arbor, MI: Frederick D. Patterson Research Institute of the College Fund/UNCF.

Newcomer, P., & Hammill, D. (1977). *The Test of Language Development.* Austin, TX: Pro-Ed.

Newcomer, P., & Hammill, D. (1988). *Test of Language Development–2: Primary.* Austin, TX: Pro-Ed.

Nippold, M. A. (1988). *Later language development: Ages nine through thirteen.* Boston: College Hill Press.

Nordenstam, K. (1992). Male and female conversational style. *International Journal of the Sociology of Language, 94,* 75–98.

Palmer, D. J., Olivarez, A., Willson, V. L., & Fordyce, T. (1989). Ethnicity and language dominance: Influence on the prediction of achievement based on intelligence test scores in nonreferred and referred samples. *Learning Disabilities Quarterly, 12,* 261–274.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black-white test score gap widen after children enter school? In C. Jenkins & M. Phillips (Eds.), *The Black-White test score gap.* Washington, DC: Brookings.

Prutting, C., Gallagher, T., & Mulac, A. (1975). Imitation: A closer look. *Journal of Speech and Hearing Disorders, 41,* 412–422.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39,* 1239–1257.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38,* 850–863.

Rondal, J. A., Ghiotto, M., Bredart, S., & Bachelet, J.-F. (1987). Age-related, reliability and grammatical validity of measures of utterance length. *Journal of Child Language*, *14,* 433–446.

Scarborough, H., Wyckoff, J., & Davidson, R. (1986). A reconsideration of the relation between age and mean length of utterance. *Journal of Speech and Hearing Research, 29,* 394–399.

Scott, C. M. (1988). Spoken and written syntax. In M. Nippold (Ed.), *Later language development: Ages nine through thirteen* (pp. 49–96). Boston, MA: College Hill Press.

Seymour, H., & Bland, L. (1991). A minority perspective in the diagnosis of child language disorders. *Clinics in Communication Disorders, 1,* 39–50.

Smitherman-Donaldson, G. (1977). *Talkin and testifyin: The language of black America*. Boston: Houghton Mifflin.

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

Stockman, I. J. (1986). Language acquisition in culturally diverse populations: The Black child as a case study. In O. L. Taylor (Ed.), *Nature of communication disorders in culturally and linguistically diverse populations* (pp. 117–155). San Diego, CA: College-Hill Press.

Stockman, I. J. (1997). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools, 27,* 355–366.

Stockman, I. J., & Vaughn-Cooke, F. (1989). Addressing new questions about Black children's language. In R. W. Fasold & D. Schiffrin (Eds.), *Language change and variation* (pp. 275–300). Amsterdam: John Benjamins.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 39,* 1284–1294.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brian, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40,* 1245–1260.

Washington, J., & Craig, H. (1992a). Performances of low-income, African American preschool and kindergarten children on the Peabody Picture Vocabulary Tests–Revised. *Language, Speech, and Hearing Services in Schools, 23,* 329–333.

Washington, J., & Craig, H. (1992b). Articulation test performances of low-income, African American preschoolers with communication impairments. *Language, Speech, and Hearing Services in Schools, 23,* 203–207.

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

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

Washington, J., & Craig, H. (1999). Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test–III. *Language, Speech, and Hearing Services in Schools*, *30,* 75–82.

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

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Research*, *38,* 1349–1355.

Wells, G. (1985). *Language development in the preschool years.* New York: Cambridge University Press.

Wiener, F. D., Lewnau, L. E., & Erway, E. (1983). Measuring language competency in speakers of Black American English. *Journal of Speech and Hearing Disorders, 48,* 76–84.

Willson, V. L., Nolan, R. F., Reynolds, C. R., & Kamphaus, R. W. (1989). Race and gender effects on item functioning on the Kaufman Assessment Battery for Children. *Journal of School Psychology, 27,* 289–296.

Wolfram, W. (1969). *A sociolinguistic description of Detroit Negro Speech.* Washington, DC: Center for Applied Linguistics.

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

Contact author: Holly K. Craig, PhD, University of Michigan, 1111 E. Catherine St., Ann Arbor, MI 48109-2054. E-mail: hkc@umich.edu