University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Publication of the DELV tests and beyond

NIH Working Groups on African American English (AAE)

2010

Developmental Trends for Features Contrastive between African American English and General American English

Janice E. Jackson

Barbara Zurer Pearson

Running Head: DEVELOPMENT TRENDS FOR CONTRASTIVE FEATURES

Developmental Trends for Features Contrastive between African American English and General American English

Janice E. Jackson

University of West Georgia

Barbara Zurer Pearson

University of Massachusetts Amherst

Acknowledgments: This work was funded in part by National Institutes of Health (NIDCD) under Contract #N01 DC8-2104 to Harry Seymour, Principal Investigator at the University of Massachusetts Amherst, with Thomas Roeper and Jill de Villiers at UMass and Smith College as co-investigators. It was accomplished in conjunction with The Psychological Corporation of Harcourt Assessments, Inc. San Antonio, TX, Lois Ciolli, Senior Research Director. Funding was also provided by the College of Education, University of West Georgia.

Correspondence concerning this paper may be addressed to Janice Jackson, Dept. of Collaborative Support and Intervention, Speech Language Pathology Program, University of West Georgia, Carrollton GA, 30118. email address: janicej@westga.edu

Abstract

Purpose: In order to provide developmental guidelines for predicted dialect differences, trends in the use of features contrastive between General American English (GAE) and African-American English (AAE) were explored in a representative, national sample. **Method**: Participants were 1130 AAE- and 194 GAE-speaking typically-developing children, ages 4 to 12. A core set of distinguishing developmental features were identified and used in a focused elicitation of contrastive items in obligatory contexts. Responses were analyzed for age-related changes and demographic influences. **Results**: Major findings include pervasive use of zero-marked forms and predicted substitutions in young typically-developing AAE speakers through age 6. After age 6 the pattern was reversed: levels of zero-marking were lower than overt-marking, but remained distinct from GAE levels of similar forms. Moreover, total absence of overt-marking after age 7 was *a*-typical. Increased use of overt-marking and other GAE-like forms was observed earlier for constructions like multiple negation and regular past tense */–ed/* than invariant subject-verb agreement (e.g., *she run*). Parent education and region, but not gender, were significant, but small factors influencing contrastive feature use. **Conclusion**: Expected contrastive feature shifts were distinguished from general language development and possible clinical indicators.

Keywords: African American English, dialect, developmental, contrastive features,

Developmental Trends for Features Contrastive Between

African American English and General American English

The identification of language impairment in children is predicated on knowing what is typical. Abundant, well-established data exist for typical developmental trends in many mainstream varieties of languages; however, relatively little comprehensive research exists for non-mainstream varieties within languages (Oetting & McDonald, 2001; Tager-Flusberg & Cooper, 1999). The lack of developmental trends for typical development in non-mainstream varieties such as African American English (AAE) has hindered researchers' and clinicians' ability to describe language impairment in speakers of those varieties (Stockman, 2010).

Studies on the acquisition of syntactic constructions specific to AAE, like negative concord (Coles-White, de Villiers, & Roeper, 2004), aspectual *be* (Jackson, 1998; Jackson & Green, 2005), and invariant *be* versus *bIn* (Green & Roeper, 2007), underline the fact that AAE is not a pidgin version of General American English (GAE), but a complex system in its own right that develops in stages. From an educational point of view, however, teachers and clinicians focus less on AAE's distinctive syntax than on typical AAE morphosyntactic and phonological constructions that overlap with analogous features of GAE often considered to be clinical markers of language impairment in that variety. These overlapping elements, in particular, are thought to impede children's progress through an academic system that requires GAE (Charity, Scarborough & Griffin, 2004). Therefore, in the clinical literature, most attention has been focused on the AAE-speaking child's transition away from usage patterns shared by typical AAE and *a*-typical GAE speakers.

Moreover, while it is possible to distinguish typical (developmental) from a-typical patterns in GAE, the absence of established developmental trends for AAE, means typical and a-typical patterns are largely ambiguous in child AAE speakers. The developmental trajectories for these ambiguous constructions differ in the two varieties in ways that are as yet poorly understood (Jackson & Pearson, 2010; Oetting & McDonald, 2001; Stockman, 2010). Many of the contrasting morphosyntactic and phonological elements that differentiate typical AAE from typical GAE are *variable use forms.* They occur in both AAE and GAE, but have different rules for their use in each variety. Most notably, variable use forms are characterized by morphological and phonological zero-marking, tense-agreement and negation variations, as well as predicted lexical and phonological substitutions. For example, most morphological endings in GAE are *overtly marked* obligatorily, while in AAE, usage varies between overt and *zero marking* of the same forms in similar contexts. So, a regular past tense construction in AAE can be either overtly marked (*yesterday Jan play<u>ed</u>*) or zero marked (*yesterday Jan play∅*), but in GAE must always be marked (*Jan play<u>ed</u>*). In the phonology of adult GAE speakers, the substitution of [f] for /θ/ is associated with immaturity (e.g. *maʊf* for *mou*θ), whereas both forms are accepted variants in adult AAE (Wolfram, 1991).

Variable use also distinguishes the varieties' agreement and negation systems, although, in this instance without zero-marking. Instead, variability is found with variable application of rules governing agreement and negation. For example, person and number dictate agreement in GAE as it employs an irregular agreement system that then determines the form of the lexical item used. By contrast, in the agreement system of AAE, usage varies between regularized forms and irregular or forms, often called invariant subject-verb agreement (Green, 2002). So, the sentence, *They were tired,* is compatible with both AAE and GAE, while in the sentence, *They was tired,* the different form of the verb *to be*, is compatible only with AAE 3rd-plural agreement. For negation, we also see the variable application of rules. While AAE allows for negation in the same manner as GAE (*He doesn't have any shoes)*, it also permits multiple negative elements in the same sentence, sometimes with alternate lexical items not compatible with GAE (*He ain't got no shoes; He don't have no shoes).* A list of the variable use forms explored in this study is found in the appendix.

As noted, variable use forms (e.g., zero-marking, regularized agreement, or variations in negation) are not consistent with mature GAE morphosyntax. Therefore, past well-established developmental points, continued zero-marking and other variants in a GAE speaker may indicate a possible language difficulty. Similar alternations between zero and overt marking, and other variant forms in AAE speakers at the same ages may also be found in AAE children with a language disorder, but are equally likely to be normal dialectal variation (Jackson & Pearson, 2008; Seymour, Bland-Stewart, & Green, 1998). Likewise, salient phonological structures and phonotactic constraints of AAE, such as substitution of /f/ for /θ/ or infrequent use of complex codas, are characteristic of adult AAE, but are rarely seen in the speech of GAE speakers after a certain age (Pearson, Velleman, Bryant, & Charko, 2009; Wolfram 1991).

Historically, clinicians working with AAE speakers have tried to balance the call for dialect sensitivity with the need to identify and treat language disorder as early as possible. The overlapping developmental and dialectal nature of zero marking and other characteristic substitutions, combined with diagnostic instruments normed on GAE, has made this task difficult (Seymour et al., 1998; Stockman, 2010). Accurate diagnosis is made even more challenging because variable usages, such as alternation between zero and overt marking and other substitutions, are known to differ not just *between* varieties, but *within* each variety in different circumstances. Therefore, typical levels of variable structures for a 5-year-old AAE speaker in Louisiana, for example, might be higher than levels seen in a same-aged AAE-speaking child in California. At present, there is no way to estimate if both are within typical ranges for AAE speakers. In the absence of established developmental guidelines for variable use features, the ambiguity of the difference/disorder conundrum remains (Seymour et al., 1998).

As with mainstream varieties, the establishment of typical developmental trends in AAE is requisite to distinguishing typical variation from a-typical performance. This paper takes a step toward identifying the developmental trends in variable use patterns based on a representative, nationwide sample of AAE-speaking children. By delineating typical changes at different ages in the most salient variable use patterns contrastive between AAE and GAE varieties, we attempt to provide developmental benchmarks for predicted variation in typical—non-disordered—AAE.

Past Studies of Contrastive Variable Use Features in Children

Although many important studies have examined variable feature use in African American children, their methods and participant samples have been as varied as their research aims. We discuss here how these data have informed us as well as their methodological and research limitations relevant to this study.

Methodological variations. Most previous studies have sampled the frequency of variable use forms like those in the appendix within relatively unconstrained free-play situations (Newkirk, Stockman, Guillory & Seibert, 2003; Oetting & McDonald, 2001; Washington & Craig, 1994), or combined with picture prompts (Craig & Washington, 2006). Although progress has been made in protocols for consistent observations from language samples (Stockman, 1996), samples show only what children *do* say, not what they *can* say, nor what they would choose in an obligatory context. However, obligatory contexts for many constructions simply do not arise with enough regularity to be able to make strong normative statements based on their possible occurrence in a language sample context.

Other studies attempt to remedy the weaknesses of language samples by providing specific opportunities for AAE forms—in reading passages out loud (Craig, Thompson, Washington, & Potter, 2003), sentence imitation (Charity, 2007), using portions of standardized tests for elicitation (Bountress, 1983; Haynes & Moran, 1989), or even a version of the classic "wug" test (Berko, 1958; Ramer & Rees, 1973). With these methodologies, there is often fuller information about the variables observed, but conclusions remain limited by the range of features examined. In Bountress (1983), for example, zero marked copula was the most persistent feature, but that study did not look specifically at invariant subject-verb agreement, which is equally persistent according to several studies by Craig and Washington (e.g. 1994, 2004). Therefore, we do not know which feature would have been more persistent for the Bountress participants.

Distinguishing characteristic features of AAE. Lists of features, considered to discriminate between GAE, AAE and other English dialects in typical populations, attempt to balance the need to be comprehensive with the need to be efficient and consistent when distinguishing between varieties. Inventories of various sizes have been proposed. Craig and Washington (2006), for example, used a list of 24 morphosyntax and 9 phonological features and various composites made from them, but in most of their charts only between seven and 10 features were used by a significant portion of the AAE speaking participants. In their effort to characterize differences between Southern AAE and Southern White English, a variant of GAE considered to be closest to AAE, Oetting and McDonald (2001) presented 35 candidate features, but these authors were able to reduce the list to ten patterns without significant loss of discrimination between dialect groups.

More recently, Wolfram and Terry and colleagues (Renn & Terry, 2009; von Hofwegen and Wolfram, in press) have taken steps to reduce the number of features examined in order to expand the kinds of analysis that can be done with data of this type. As Renn and Terry point out, statistical options are limited when there are a large number of variables under examination. Therefore, they performed a rigorous comparison of the ability of measures of different sizes to characterize differences in usage patterns for the same participants in two contexts that differed greatly in formality. They used the Craig and Washington (2006) inventory as a benchmark, but also added other features to it in case they would be more informative. Then they also reduced the list of candidate features drastically to a subset of six features. Near-perfect correlations between the three measures showed that adding features to Craig and Washington's inventory did not increase explanatory power and taking away features reduced it only slightly. They concluded that a judicious selection of features will greatly expand statistical options for large-scale studies. Similarly high correlations found by Von Hofwegen and Wolfram (in press) among alternate inventories validate the subset strategy and suggest that the ultimate combination of features permits some variation.

Summary measures of contrastive feature use. In order to facilitate comparisons across individuals and groups, different types of summary measures have been proposed: 1) token-based "dialect density" measures; 2) type-based summaries of contrastive features, and 3) variation analysis based on a subset of types from type-based summaries (von Hofwegen & Wolfram, in press). Token-based density measures typically consider the number of forms that diverge from GAE analogous forms as some proportion of total utterances or words (Oetting & McDonald, 2002). In this way, variations in the number of features zero and overtly marked, for example, are conceptualized as higher or lower levels of dialect density. Alternately, researchers report either the number of types used by individuals, or the percentage of individuals who used a certain feature type (Craig & Washington, 2004; Jackson & Pearson 2010; Washington & Craig, 1994). In many of these studies, descriptions of high, medium, and low levels are made from a particular corpus using cluster analysis after the fact, and thus they may not generalize to other corpora.

Overall, the objective has been to find for AAE what already exists in the developmental data for mainstream varieties; that is, a core set of morphosyntactic and phonological features that reliably predicts typical performance across multiple age ranges. Even before looking at comparison groups of language-impaired children, characterizing typically developing children will indicate a baseline against which other comparisons can be made.

Past Studies Capturing AAE Within-group Variability

Whether conceptualized as dialect density in terms of tokens, types, or opportunities, the tendency to minimize variable use patterns in contexts that encourage GAE use has been shown to respond to many factors. Such factors may cause the identification of what is “typical” to vary within subgroups of speakers with different backgrounds, and so have the potential to require different guidelines for different subgroups. In discussions of AAE, the most often reported among such factors, in addition to age, are parent education level and/or community *socioeconomic status* (SES) (Craig et al., 2003; Horton-Ikard & Miller, 2004; Seymour, Roeper, & de Villiers, 2005); *gender* (Craig & Washington, 2002; Ratusnik & Koenigsknecht, 1975; Washington & Craig, 1998); and *region* (Charity, 2007; Newkirk et al., 2003).

Age. Clearly age is the key factor for discussions of development, but greater changes in variable feature use appear to occur at some ages more than others. Many previous studies show that in school and clinical contexts, the use of contrastive features like those in the appendix declines over time, at least during elementary school (Charity, 2007; Craig et al., 2003; Craig & Washington, 2004; Ratusnik & Koenigsknecht, 1975; Von Hofwegen & Wolfram [in press]; Washington & Craig, 1994, among others), but the timing differs across studies. Both Craig and Washington (2004) and Charity (2007) found a significant break between kindergarten and first grade for morphosyntax features, but Charity found a similar break for phonology features as well, whereas Craig and Washington did not probe phonology until later ages. In Haynes & Moran (1989), a significant increase in the number of final consonants pronounced occurred between pre-kindergarten and kindergarten. In Isaacs (1996), a similar shift came between third and fifth grades.

Furthermore, not all changes in variable feature use can be attributed to development solely within the particular variety (AAE). Several authors have noted that many of the features of interest for AAE speakers are undergoing developmental changes in mainstream populations as well. Isaacs (1996) observed changes similar to those in her AAE participants in a comparison group of GAE speakers. Craig, Thompson, Washington and Potter (2003) also acknowledge the influence of general developmental features in their decision not to analyze phonological features of the dialect in children younger than second grade. So, discussions of age-related factors in AAE need to also make reference to general developmental trends, unrelated to dialect.

SES. Another factor that might influence how developmental trends are framed is socio-economic status (SES). The literature presents an inconsistent picture of the effects of SES on contrastive feature use. Some studies identified SES as a main effect (Ratusnik & Koenigsknecht, 1975; Washington & Craig, 1998), while others tested for it and found no significant effect (Craig & Washington, 2002; Thompson, Craig & Washington, 2004). The typical pattern is for higher SES to be associated with lower levels of dialect density (Ratusnik & Koenigsknecht, 1975; Washington & Craig, 1998). But the relatively few studies of mid-SES African American (AA) children (Horton-Ikard & Miller, 2004, Ogbu, 2003) showed widespread use of AAE features in that population as well. In all populations, use of forms comparable to GAE have been shown to respond to differences in both setting and topic (DeBose, 1992).

Variable use patterns typical of AAE also appear to vary by SES for different features. In Bountress (1983), for example, certain phonological features followed the expected pattern of more overt marking among higher SES participants, but the zero marked copula was relatively resistant to change among the AAE speakers at both higher and lower SES schools. As pointed out by Horton-Ikard and Miller (2004), SES may moderate the effect of dialect density, but child density levels overall appear relatively high at all levels of parental SES.

Gender. Gender may also be a strong enough influence on dialect density that separate benchmarks would be needed for girls and boys. Perhaps because of early reports on AA adolescent males (Labov, 1970; Wolfram, 1969), boys are sometimes thought to be heavier users of dialect features than girls. Few studies, however, support such a claim for young children. A small number of studies did indeed find significant differences favoring more contrastive feature use by younger males (Craig & Washington, 2002; Washington & Craig, 1998). More common, however, is the situation observed by Charity (2007), Horton-Ikard and Miller (2004), and Washington, Craig, and Kushmaul (1998) where the gender difference was evident in some circumstances and for some features but not others. When gender effects have been found in these studies, they have not always been in the same direction and have been hard to interpret.

Region. Historically, AAE has been unusual in that it has not been considered a geographic dialect (Labov, 1970), although more recent studies, like Pollock & Berni (1996), have pointed to regional differences, at least in phonology. There are relatively few studies that address the issue of region of residence directly by comparing more than one region within a single study. Newkirk et al. (2003) compared children in Michigan with Louisiana, and Charity (2007) worked in Ohio, DC, and Louisiana. Both authors found that the types of features used were similar in the North and the South, but the frequency of tokens was greater in the South. However, the Newkirk et al. (2003) study involved only ten preschoolers in each location. There were more participants in Charity's study and though the effects were statistically significant, she reported only moderate effect sizes for region.

The Current Study

The review of previous research leads to the following considerations for the current study. Critical to identification of developmental trends is the examination of a large number of children independently identified as typically developing. Also useful, would be the inclusion of a comparison group of children learning GAE as a first dialect in order to tease out the effects of development more generally. Finally, the participant sample would need to be comprised of larger numbers of speakers from more diverse socioeconomic backgrounds and geographical areas than have previously been available in order for findings to be readily generalized. These criteria suggest a structured elicitation procedure to provide an obligatory context for zero or overt marking and other structures, so the elicitation can be more uniform and more informative for larger numbers of participants. Further, using suggestions from the literature as a guide, a relatively extensive selection of the most promising characteristic features from past studies should be employed first and then carefully narrowed based on statistically identified trends. Developmental trends identified in this manner will go a long way to help educators and clinicians rule out predicted dialect differences as patterns of disorder in the language of AAE speaking children in GAE academic settings.

Therefore, for the present study using data from elicitations performed on two extensive samples of children, answers to the following questions were explored:

Research Question 1) From the set of contrastive variable use features that distinguish AAE and GAE child speakers, which core set of features has the potential to give the most information about AAE developmental trends?

Research Question 2) What general developmental trends in AAE can be identified from a core set of variable use features (Research Question 1) and how are they modified by demographic factors, specifically SES, gender, and region?

Method

Participants

Two nationwide samples of children participated in this cross-sectional study that involved, in all, 1,130 African American (AA) and 194 Euro-American (EurA) children in nine age groups from 4 to 12 years. For clarity, participant groups are described both in terms of relevance to the research questions and specific participant group characteristics.

Sample 1 (Research Question 1). Both AAE and GAE participants comprised this sample.

AAE participants. AAE participants for Research Question 1 were 443 typically-developing AA children ages 4 through 12, who were part of the field-testing sample for the *Diagnostic Evaluation of Language Variation* (*DELV) Norm Referenced (NR)* and *Screening Test (ST)* (Seymour, Roeper, & de Villiers, 2005, 2003), initially named the *Dialect Sensitive Language Test* (*DSLT*) during the field-testing phase. Participants were recruited from all regions of the country, and included 55% females. In order to maximize the number of variable use forms and AAE patterns observed, recruitment followed suggestions in the literature reported earlier, and so the sample had a high percentage of low-SES children and an over-representation from the South.

Specific criteria for membership in the AAE group were 1) African American background (by parent self-report); and 2) residence in a community of predominantly African Americans by census report (U.S. Bureau of the Census, 2000). 3) The third criterion was that participants be typically developing according to standard speech-language assessment practices; that is, they were recruited by certified speech-language pathologists (SLPs) as typically developing, and none were receiving speech-language services. In addition, they scored in the passing range on the subset of *DSLT* items that eventually came to comprise the *DELV-NR* (i.e. within 1 standard deviation of the mean or higher).[1] All participants also demonstrated average or above performance on a subset of *non*-contrastive morphosyntax items from the *DELV-ST* Part 2. No selection was made on dialect density. Among the children who met the three selection criteria, there was a range of dialect densities from *no difference from GAE* to *a strong difference from GAE*.

GAE first-dialect participants. A comparison group of 194 typically-developing EurA GAE speakers, also from the field-testing, were recruited from predominantly EurA communities by census report, and were matched to the AAE participants on age, region, gender, parent education, and clinical status. Like the AA children, they were 55% female, primarily low SES, and from all regions of the U.S. None were receiving speech-language services in their schools, nor were any of them below average on the *DELV-NR* or the noncontrastive morphosyntax in the *DELV-ST* Part 2.

[1] There were no scores on the *DSLT*, which these participants took. The *DELV-NR* norms, based on another participant sample, were used later when they became available to confirm typical development.

The numbers and demographic characteristics of Sample 1 are shown in Table 1.

Put Table 1 about here.

Sample 2 (Research Question 2). For Research Question 2, to permit broader generalization of our findings, a second, larger, more balanced sample was recruited based on census information about parent education level and region of residence for the general U.S. AA population (U.S. Bureau of the Census, 2000). Six hundred eighty-seven children from the preliminary National Institutes of Health (NIH) AA norming sample for the *DELV-NR* comprised Sample 2. They had been identified as typically developing by the participating SLPs, and none were receiving speech-language services in their schools or communities. Further, all participants scored within one standard deviation of the mean or higher on the *DELV-NR*, and were average or above on the subset of non-contrastive morphosyntax items from the *DELV-ST* Part 2.[2] These participants, too, represented a range of dialect densities. As shown in Table 2, they closely matched the census distributions.

Put Table 2 about here.

Combined AAE sample (Research Question 2). Participants in the combined sample were the 687 AA children in the preliminary NIH norming sample for the *DELV-NR* and the 364 4- to 9-year-old AA children in the *DSLT* field-testing sample, i.e., Sample 1 (up to age 9) and

[2] Note that after the AA standardization for NIH, there was a subsequent standardization of the *DELV-NR* based on a norming sample that represented the general U.S. population, including all ethnicities. The published norms, which were used in this study once they were available, were produced from the second norming sample, not the preliminary NIH AA sample in the current study.

Sample 2 combined. Participants in the combined sample, therefore, were 1051 children across the various categories shown in Table 3. Rationale for the combined participant sample is discussed in the analysis section in Procedures.

Put Table 3 about here.

Materials

Stimuli for identifying a core set of items (Research Question 1). The contrastive portions of the unpublished *DSLT* (Seymour, Roeper, & de Villiers, 2000) were used as stimuli. The *DSLT* was comprised of 350 items organized in 14 subtests, encompassing mostly *non*-contrastive syntax, pragmatics, semantics, morphosyntax, and phonology. While largely non-contrastive, the *DSLT* also included a less extensive set of contrastive items. Thirty-four contrastive morphosyntax items testing 11 contrastive structures from the 53-item Subtest 1 were used, as well as 55 phonology items, testing 37 phonemes and consonant clusters, from the 132-item Subtest 13. These 34 plus 55, or 89, items were utilized to identify the set of core features for Research Question 2.

All 89 items were production items aimed at eliciting specific contrastive responses from the children. The morphosyntax stimuli had picture prompts to aid in memory and interpretation. For example, one 3rd –*s* item showed two pictures side by side, one of two children riding horses and one of a single child riding a bike. The examiner pointed to the pictures in turn, saying, "I see horses. I see a bike. [pause] The girls always ride horses, [pause and point to the boy], but the boy always…." Phonology items used pictures to elicit production of targets embedded in full sentences that the child was asked to repeat after the examiner (e.g., 'I see that fish brea<u>the</u>

underwater'). The 55 contrastive phonology targets included 20 different final consonants and 17 final clusters, plus interdental fricative /ð/ in initial and medial, as well as final position (Pearson et al., 2009). The record form listed the target word with the target segments highlighted, along with the International Phonetic Alphabet (IPA) spelling of the target with space for examiners to check matching responses or note mismatched ones. The structures and number of items for each are listed in the appendix.

Stimuli for demographic analysis (Research Question 2). The 15 most contrastive items identified from Research Question 1 were used, 10 morphosyntax and 5 phonology. The elements in the core set are noted in the appendix. Although these are the 15 items that later became the dialect density identification section of the *DELV-ST* (Part 1), in the present study, actual responses, not *DELV-ST* coding categories, were used.

Procedures

Question 1, Step 1: Core set selection and dialect group trends. ***Data collection.*** Children in Sample 1 were administered the 14 subtests of the *DSLT* (Seymour et al., 2000) all in the same order. The morphosyntax items were in the first subtest, and the phonology items were given next to last. The test was administered individually by a certified SLP in a quiet room at the child's school or clinic. Younger children typically did the test in two sessions on different days. Over 400 SLPs participated in administration, almost all of whom were of European-American background, reflecting the ethnic composition of the profession. Portions of 60 *DSLT* administrations were audio-recorded for reliability.

Data analysis. Analysis of Variance (ANOVA) was performed with the percentage of AAE-like responses for each item type and then for all 89 tokens together as the dependent variable. Independent variables for Research Question 1 were dialect group (AAE- and GAE-speaking) and age categories in six levels from 4 to 12 years, collapsing two years for ages 7 to 12 years, where the numbers of participants in each year were smaller.

Correlations were done to identify a subset, or "core," of items with the highest correlation to the full set of 89 items with respect to variable use responses.

Question 1, Step 2: Establishing trends for low and high levels of variable use forms.

Data collection. Responses to the 15 items selected as the core set in Step 1 were extracted from the response database for Sample 1 participants.

Data analysis. The percentage of AAE-like variable use responses to the 15-item core set was summarized for each participant (AAE and GAE) in a Dialect Density Ratio (DDR) as follows: The number of AAE-like responses to the 15 items was divided by the total number of AAE- and GAE-like responses (DDR = AAE/ (AAE + GAE)), excluding the responses marked "other." "Other" responses were ones that gave no information about the child's dialect status. For example, if the target was a third-person singular form of *don't* (AAE-like) versus *doesn't* (GAE-like) and the child responded *didn't,* which is the same for both dialect groups, that item was not included in the denominator.

Only the EurA GAE-speakers' DDRs were used to establish a preliminary benchmark for low and high levels of variable usages for each age. A DDR that was within one standard *error* of the GAE speakers' mean for that age was considered to be "not different from GAE." DDRs that were higher than 2 standard deviations above the GAE children's levels of variable usage were considered to show a "strong difference from GAE."

Question 2 (Demographic analysis). ***Data collection.*** The 15-item core set selected from the *DSLT* items in the analysis for Question 1 as the most discriminating between AAE and GAE speakers were administered to Sample 2. In addition, for purposes of reliability of the dialect density measures derived from the core battery, language samples were collected from 78 children within a week of the administration of the test.

Data analysis. A DDR was calculated for participants in Sample 2. First the DDRs from Sample 2 alone were averaged by age, and compared to the developmental trends established from the DDRs of the EurA participants in the previous step.

Finally, DDRs for the combined samples were the dependent variable in an ANOVA. Independent variables were region (Northeast, North Central, South, and West), age (six levels from 4 to 9 years), gender, and parent educational level (PED), counting the four levels noted in Tables 1 to 3, which is a common proxy measure for socio-economic status (SES). As one can see in Table 3, despite the great number of children involved, not all cells were large enough to do an ANOVA with all of these variables, so preliminary tests, described in the Results, were performed to see which levels of the variables could be collapsed or eliminated with minimal loss of generalizability.

Reliability

Research Question 1 reliability. Reliability procedures were done for the *DSLT* as a whole. The research group had access to 60 audiotapes of selected sections of the *DSLT* testing sessions. Tapes containing the phonology subtest were transcribed and scored independently by a project consultant, who was both a linguist and certified SLP, and by the second author, also a linguist. Their consensual transcriptions of the 410 target tokens (including each element of a cluster separately) showed 94% agreement with the examiners' notations. Morphosyntax scoring involved just checking off the child's answers on a pre-printed form, so no tapes were made of the morphosyntax section.

Research Question 2 reliability. To confirm the concordance of traditional dialect measures with dialect density categories ("no difference from GAE" and "strong difference from GAE"), which were the basis for the DDR, 78 language samples were transcribed under the supervision of the second author by graduate students in Communication Disorders trained in recognizing AAE features. Each tape was transcribed and coded by two researchers, and disagreements were resolved by consensus with a third party, an AAE speaker who was an advanced doctoral candidate. The coding assigned by the *DELV-ST* Part 1 was compared to a tally of AAE features in the language samples following the Oetting and McDonald (2002) "Token 3" method for deriving dialect density. Category agreement analysis revealed that 71 of 78 children (or 91%) had dialect density scores from the language sample that were consistent with the dialect density measure in this study (Pearson, de Villiers, Magaziner, & Sunderland, 2005, and under review).

Summary of Procedures

Before turning to the results, let us summarize the steps in the analysis. Recall that the goal of Research Question 1 was to establish core patterns of variable contrastive feature use in AAE and GAE speakers at different ages. To accomplish this, Sample 1 participant responses to 89 contrastive items on the *DSLT* were used to establish patterns of zero-marking and variable usages in AAE and GAE speakers between ages 4 and 12 years. Dialect group patterns were established for the whole subset of 89 *DSLT* items and also individually for 11 contrastive item-types. Fifteen items (tokens) were identified that best distinguished the two dialect groups across the age range, i.e. were used contrastively by the largest number of children and had the highest correlation to the larger set. Responses to just those 15 items were extracted from the Sample 1 response database and a DDR was calculated for each participant. Average DDRs for the GAE-speaking group established benchmarks at each age for levels of variable use for those features that showed “no difference from GAE” or “strong difference from GAE.”

For Research Question 2, the 15-item core battery was administered to a new, larger sample of African Americans that was carefully matched to the U.S. census (Sample 2). Response patterns to those items from the 687 participants in Sample 2 were translated into the DDR, and analyzed by age to find representative patterns of development for AA children, ages 4 to 9 years. Finally, in order to explore the effects of parent education level, region, and gender on the developmental patterns of dialect density observed with Sample 2, the two samples were combined to yield the largest available number of participants. The DDRs calculated from the 15-item core battery for the two samples combined were analyzed with an ANOVA with age, gender, region, and parent education level as independent variables.

Results

Research Question 1

Step 1: Core set selection and group trends. Although these were not longitudinal data, the battery of 89 morphosyntax and phonology items confirmed that in the AAE sample, levels of zero marking and predicted variable AAE usages were generally high among the children 4 to 6 years of age and were dramatically lower for children older than age 6. The age pattern was similar among the GAE speakers, but the level of use and magnitude of the changes was smaller for that group. As shown in Figure 1, 4-year-old AA participants used AAE zero-marking and substitutions in around 33% of opportunities in the elicitation protocol compared to 10% such responses among the GAE 4-year-olds. At ages 9 to 12 years, zero-marking and predicted variable uses were observed at levels approaching 10% of opportunities for the AAE group, while analogous usage among the GAE group was near zero.

____

Figure 1 here

____

As shown in Figure 1 and confirmed in the ANOVA below, zero marking of these features by the GAE speakers was much lower across the age span--an average of 22 percentage points lower at age 4 and approximately 10 percentage points lower from ages 9 to 12.

In a univariate ANOVA with the total number of AAE-like responses as the dependent variable, dialect group and age had large effect sizes, and there was a small interaction between them: Dialect Group, F (1, 625) = 295.1, $p < .0005$, $\eta^2 = .32$; Age, F (5, 625) = 40.3, $p < .0005$, $\eta^2 = .24$; and Dialect Group by Age, F (5, 625) = 7.0, $p < .0005$, $\eta^2 = .05$. The analysis of simple effects at each age level showed that the dialect group differences were significant at each age, but the difference between dialect groups was greater among the younger children than among the older children. Tukey HSD post hoc comparisons confirmed the pattern of age differences observed in the figure: ages 4 and 5 were not different from each other, but were different from all other ages. Age 6 was different from all other ages, and age 7-8 was different from ages 9-12, but ages 9-10 and 11-12 were not significantly different from each other. Significant age effects were similar for the GAE speakers, except for this group ages 7-12 did not differ.

Individual feature patterns. Not all features followed the same developmental path. Eight of the more frequent contrastive morphosyntax features and the three most frequently contrastive phonology features are shown in Figures 2, 3, and 4.

The four features in morphosyntax that showed the most change over time and the greatest persistence among the AAE-background children were the various forms of invariant subject-verb agreement (Figure 2). Figure 3 shows that multiple negation and regular past tense

/*–ed/* had the lowest frequency and the earliest tendency toward structures consistent with GAE. Zero possessive /*'s* /, also in Figure 3, was very persistent, but showed little change over the age range. Note that zero marking of *is* and *are* copula and auxiliary forms, also in Figure 3, was not elicited at high levels with our materials.

For phonology, although there were a large number of elicitation items, in this formal test setting using single sentences, 95% of the targets were pronounced in a GAE-equivalent manner by typically developing AAE-speaking children of this age group. (See also Pearson et al. 2009; Velleman & Pearson, 2010.) Shown in Figure 4 are the most persistently contrastive phonological elements: substitutions in final position of /f/ for /θ/, /v/ for /ð/ and /f/ for /-ft/.

Figures 2, 3, and 4

***Relationships among measures*.** In order to establish which core subset of features could be used reliably in place of the broader set from Question 1, correlations between levels of variable usages for each of the individual features and the total variable usage score (for the 89 items) were obtained. The elements in Figure 3—past tense /-*ed*/, multiple negation, zero-copula/auxiliary, and possessive /-*s* /--contributed only moderately to the variability in the total variable usage score: r (N=443) = 0.49, 0.52, 0.6, and 0.36, respectively, p = .001. The correlation for the three most persistent phonology item types, shown in Figure 4, had a higher correlation with the total score at $r = 0.75$ ($p < .0005$). An even larger correlation with the total score of 0.88, $p < .0005$ was found between the four morphosyntax item-types in Figure 2, considered together, and the total score. When the 3rd-*s* verbs and the phonology items were combined, the correlation of scores for those 22 items with the scores for the 89 items rose to

0.93, *p* < .0005. So these seven item-types were chosen to comprise the core subset of features in the shorter battery, selecting 15 items: four tokens of 3rd -*s* with lexical verb*s*, one final consonant cluster, and two tokens each of the other five types: /f/ for /θ/, /v/ for /δ/, *have/has, do/does, and was/were* alternations. Those 15 items also showed a significant 0.92 correlation with the 89-item battery. Thus, they were adopted for Research Question 2 as an efficient proxy for the larger item set.

Step 2: Establishing trends for low and high levels of variable use forms. To provide a quantitative measure of dialect density, the Dialect Density Ratio (DDR) was calculated from each participant's pattern of responses to the 15-item core battery, as described in Methods. Thus with these 15 items based on the 7 item-types, *low* and *high* levels of zero-marking and other contrastive responses were established from the EurA participants' mean DDRs, providing a range of typical developmental variation independent of AAE dialect. These trends are presented below under Research Question 2.

Research Question 2

Overall developmental trends (Sample 2 and Combined Sample). Figure 5 shows the DDR median and one standard deviation above and below the mean at each age between 4 and 9 years for the 687 typically developing children in Sample 2. The levels of "no difference from GAE" and "strong difference from GAE" for each age established above are noted on the figure.

____

Figure 5 around here

____

The use of the error bars emphasizes the fact that at every age, but especially at ages 4, 5, and 6 years, very high levels of AAE patterns of response were more common than low levels among these typically developing children. Among 4-year-olds, only 2% of the AAE children used no zero-marking in response to the stimuli, and only 10% used levels comparable to "no difference from GAE." Fully 26% of these typically developing 4-year-olds responded with 15 of 15 zero-marked responses or substitutions. The AAE median DDR at age 4 was calculated to be 0.8. Fifty percent of these typically developing children had DDRs of 0.8 or higher--roughly equivalent to 12 of 15 zero-marked responses, and 0.2 points higher than the "strong difference from GAE" benchmark, also shown in the figure.

These proportions reversed for children older than 7 of whom only 2% had DDR > 0.8 and 17% used no zero-marking at all in the elicitation. Still, even though exclusive use of zero marking after age 7 was very rare, DDRs for the older children showed that they used zero-marking and predicted substitutions for these stimuli at rates significantly higher than GAE-speaking children. As seen in Figure 5, rates of zero-marking up to DDR = 0.6 (that is, roughly 9 of 15 zero-marked responses) were within one standard deviation of the mean throughout the age range, and 75% were above the "no difference from GAE" benchmark established by the EurA participants. The median values of DDR at each age show that a 0.4 DDR on this measure was very low for a 4- or 5-year-old but high for a 9-year-old. Similarly, 0.8 DDR was average for a 4-year-old and extremely rare for a 7- to 9-year-old.

Correlations between separate morphosyntax and phonology dialect density ratios and the combined DDR confirmed that both morphosyntax and phonology contributed significantly to the total contrastive feature ratio. For the AAE speakers, the DDRs for morphosyntax and phonology were correlated with each other (r [N= 443] = .53, $p < .0005$) and the components were correlated at 0.93 and 0.81 respectively with the total, also significant at $p < .0005$. Since there were more morphosyntax items, it is not surprising that the correlation to the total score was higher than for phonology. Nevertheless, the significant correlation for the phonology DDR shows that those items, too, contributed to the effects observed.

Demographic trends. ***Relative strength of the four major demographic factors at different ages*****.** To increase power for the test of the demographic variables for Research Question 2, Sample 1 and Sample 2 were combined as indicated in Table 3. Table 4 reports DDR values for different subgroups of the combined sample (N = 1051) according to the variables targeted in Research Question 2: age, region, parent education level and gender. As one can see from the paired superscripts in the table, only four comparison pairs fell within the 95% confidence intervals for each other: the DDR for ages 7 and 8, ages 8 and 9, PED levels 1 and 2, and gender. Means for the other levels of the variables were all quite distinct.

Put Table 4 about here

An ANOVA with DDR as the dependent variable and age, PED level, region, and gender as independent variables was planned. However, even with over 1,000 children in the combined sample, not all 192 cells of a 6 x 4 x 4 x 2 ANOVA would be of sufficient size. Therefore, univariate analyses with each variable separately helped reduce the number of variables for the full ANOVA. For example, as one might predict from the DDR values for gender in Table 4, gender was not a significant main effect: F (1, 1049) < 0.1, p = .4, η^2 = .001. As a check on potential interactions, gender was entered in separate ANOVAs with each of the other variables: Gender by Age: F (5, 1039) = 1.2, p = .29, η^2 =.006; Gender by PED level: F (3, 1043) < 1, p = .5, η^2 =.002, and Gender by Region: F (3, 1043) < 1, p = . 9, η^2 <.0005. Since there was no main effect nor any significant interactions, gender was not entered in the full ANOVA.

Next region was considered. Regions were represented in the combined sample in approximately the distribution of the U.S. AA population (as shown in the rightmost column of Table 3). Thus, the number of participants from the south was very large and from the west was quite small, with the other two regions in between. A univariate ANOVA showed a main effect of region (F (3, 1047) = 19.24, p < .0005, η^2 = .05) and in pairwise comparisons, each region was significantly different from the others. However, the west especially lacked sufficient numbers to split by age and parent education level. Even if levels of age and PED were further collapsed (as indicated below), some cells for the west would still have fewer than 10 participants. So, the children from the west were not entered into the full ANOVA, leaving an N of 978.

Turning to the continuous variables—age and parent education level--it was noted in Table 4 that not all of the age groups were outside the 95% confidence interval for all other ages with respect to DDR. In Tukey-B tests, ages 4 and 5 years were one category (p = 0.05), and age 8 was not different from 9, so those four ages were collapsed into two. Age 7 was not different from age 8, but it was different from 6 and 9, and so it was maintained as a separate group, leaving four levels of the age variable; 4-5, 6, 7, and 8-9 years.

Finally, with all four parent education levels, several cells were still too small for meaningful analysis, especially in PED levels 1 and 4 (less than 11 years of education and a 4-year college degree, respectively). The preliminary univariate ANOVA showed a significant effect of PED level (F (3, 1047) = 36.18, p < .0005, η^2 = .09). However, as also indicated in Table 4, pairwise comparisons showed that levels 1 and 2 were not significantly different from each other, so it was an easy decision to collapse them into one category of "high school degree or less." However, level 4 was significantly different from level 3 and so a solution was not obvious. In the end, we report the ANOVA below with levels 3 and 4 collapsed, i.e. with two levels of PED, high school degree or less versus more years of schooling past high school. (The full analysis was also performed with three levels of parent education. The *F*-values changed, but the same effects were significant, even down to similar effect sizes.)

In the final 4 x 3 x 2 ANOVA, age and region were significant main effects as follows: Age: F (3, 954) = 71.40, $p < .0005$, $\eta^2 = .18$ and Region: F (2, 954) = 11.57, $p < .0005$, $\eta^2 = .02$. The main effect of age was based on the significantly lower DDRs among the older children compared to the younger children, as predicted. The effect of region was somewhat unexpected. As shown in Table 4, and in pairwise comparisons with the three regions in this analysis, the mean DDR for the North Central participants was significantly higher than the mean for those from the South or the Northeast. There were no significant interactions with either age or PED, indicating that the regional differences were consistent at all levels of those variables.

The PED level effect was larger than the effect for region: PED: F (1, 954) = 76.86, $p < .0005$, $\eta^2 = .07$. A small Age by PED level interaction (F [3, 954] = 4.32, $p = .005$, $\eta^2 = .01$) reflected the fact that although the parent education levels were significantly different at each age, the size of the effect differed at different ages. In the analysis of simple effects, the effect size of the PED level variable was large at ages 6 and 7, $\eta^2 = 0.17$ and $\eta^2 = 0.36$ respectively, but only $\eta^2 = 0.06$ at age 4-5 and $\eta^2 = 0.08$ at age 8-9.

Put Figure 6 about here

In particular, as shown in Figure 6, the age effect varied in the same direction at both levels of parent education, but the time course of the decrease in DDR was different for each parent education level. For children whose parents had a high school education or less, the DDR of 6-year-olds was not very different from 4- and 5-year-olds'. Then a DDR lower by 0.10 or more at each year was seen for the age groups 6, 7, and 8-9 years. By contrast, for the children of parents with more education, the DDR was almost 0.20 lower at age 6 than age 5, and then 0.2 lower again for age 7. For the higher parent education level, there were no significant differences in DDR observed between ages 7 and 9.

Discussion

Based on the previous literature, there was little doubt at the outset of this study that among AAE-speaking children exposed to GAE as a second dialect, AAE contrastive features would occur at lower levels in the later elementary school years than at pre-kindergarten, but it was not clear when the significant changes would occur. The effects of region, parent education level and gender were more difficult to predict because reports on these topics were less unanimous in the literature, or they came from smaller, or more localized studies. We anticipated that parent education levels would have a relatively large effect, as it does on so many other aspects of language development (Hoff, 2003), and we expected possibly significant, but smaller effects of region and gender.

In fact, like others before us, we observed lower levels of AAE variable usages and higher levels of usage comparable to GAE at age 9 compared to age 4. However, levels of variable usage for some features remained higher and more widespread longer than the literature led us to expect, especially at ages when similar developmentally related zero marking and substitutions were all but absent among GAE controls. With respect to the effect of demographic variables, our expectations were both met and not met. There were significant effects of parent education level and a smaller, but significant effect of region. The effect size for parent education was higher than for region (η^2 = 0.09 vs 0.05). Still, both were relatively small.

Furthermore, the effect of region ran counter to our expectations. It was not the South but the North Central region that had the highest levels of zero-marking and other substitutions associated with AAE, even when correcting for the lower parent education levels found in the North Central cohort compared to the other three regions. At the scale of these data, gender did not approach significance.

Previous literature, however, was less able to inform us about what levels of zero marking and variable usages could be considered developmental, or what level of contrastive feature use could be considered a-typical for school-aged AAE speakers. Should we expect older AAE children to speak completely without zero marked features in their speech, at least in a school environment? What proportion of typically developing children continue to use high levels of zero-marking through age 9, and will the patterns be the same at all levels of parent education, region, and gender?

In the past, generalizable answers to these more specific questions have been either inconsistent or unattainable given sample size and methodological constraints, as well as geographical and SES restrictions. With the use of a standardized methodology and greater demographic balance, we have attempted here to expand the information base from which these more detailed questions could be answered. We found both converging results and some new findings.

Developmental Trends

Differences from previous findings. The work presented here differs from previous studies in several key respects, sample size, geographic location, number and kind of features studied, and methodology. 1) The current study has a larger sample and is, therefore, more generalizable; 2) It has a more diverse geography and therefore may serve to better resolve earlier inconsistent reports about geographic influence. 3) A third advantage of this study is its structured elicitation that created an obligatory context for variable use forms. Thus, it overcame limitations in other methodologies where opportunities to use specific features may not have occurred and resulted in lower estimations of frequency. Elicitation protocols present the same opportunities to everyone, and so, even though rates appear higher here than in other studies, the data are more comparable across groups. Finally, 4) the formal testing setting of this study was more likely to elicit GAE-like responses by AAE speakers who were able to code-switch, while at the same time revealing the variable use patterns that persisted in academic and formal settings. A further advantage was that responses to the standardized elicitations were easily analyzed and did not require familiarity with the target variety.

One finding that stands out as different from previous findings relates to regional variation. Although region has not been considered a source of specific differences in subdialects of AAE (Labov, 1970), when there have been differences noted, the studies indicated that dialect density was greater in the south than in other parts of the country (Charity, 2007; Newkirk et al., 2003). Consistent with our general expectation, in these data, region was a significant but small effect, $\eta^2 = 0.05$[3], with children from the West having the lowest average DDR (0.24), then the Northeast (0.38), then South (0.47). Children from the North Central region (Detroit, Chicago, Indiana, etc.), had the highest average DDR (0.57). Its DDR was higher at all ages and at all PED levels. One speculation is that this difference was associated with an urban/ suburban distinction, with more concentrated urban areas with large numbers of African Americans found in the North Central. The higher DDR, even for children of families from similar levels of parent education, may stem from greater social isolation of AA

[3] This value is taken from the preliminary ANOVA with all four regions in the analysis.

communities in the north. Or, different features may experience regional influences differently. Further research will be needed to elucidate these effects.

Convergent and expanded developmental findings. Dramatic differences in the rate of zero marking among participants were found between ages 4 and 9, roughly pre-kindergarten and grade 4. One can clearly see in Figures 1, 5 and 6 a shift away from high levels of variable usage in the direction of increased levels of GAE-like forms from age group to age group among typically developing children in both samples. This pattern is consistent with general findings from Craig and Washington (2006) and VonHofwegen and Wolfram (in press), although data from the latter extend into teenage years where trends were reversed.

As documented in prior research, a higher level of overt marking and GAE-like usages was observed after age 6 than before age 6. While a small portion of the variable feature use among the AAE-speaking children was analogous to the similar use of the same features observed in younger EurA children, these features occurred in the EurA children's responses at much lower rates, and were almost at zero after age 6 when most major developmental shifts in GAE for these constructions have ceased. The EurA children, who had no known exposure to AAE, thus provided an estimate of the portion of the variable usages among the AA children that might be unrelated to their exposure to AAE i.e., would be more developmental in nature.

It has been proposed that the critical changes between ages 6 to 7 coincide with most children's entry into first grade and increased exposure to GAE literacy practices (Charity et al., 2004). However, school entry was likely not the only factor influencing changes at these ages. Parent education level was shown to affect not only the overall level of variable usage, which was higher for families with less education, but also the age intervals at which the decreased levels of variable usages were observed: as illustrated in Figure 6, differences in variable usage levels between ages 5 and 7 were very large for the higher level of parent education, whereas for those with lower levels of parental education large differences were observed only after age 6, between 6 and 8 years of age.

Examining AAE-speaking children alongside a GAE control group, one cannot easily determine for either variety what portion of this shift is attributable to school entry and what part to maturation. Indeed, it cannot be assumed that all AAE speaking children enter a school environment where GAE is the only variety spoken. It is possible that in certain communities, perhaps those with strong cultural and linguistic ties, AAE speaking children are served by AAE speaking teachers and staff, in which case the GAE impact would not be the same. Moreover, as all typical children move toward more adult-like use of their target variety, one would expect there to be developmental shifts (similar to this one) at various points along the way. However, in the case of AAE, these points are as yet poorly understood especially in terms of their relationship with typical variation in adult AAE. In addition, the task to determine which socio-cultural and linguistic variables lend themselves to increased awareness and/or pressure to generate code- or style-switching in child speakers, compounds the difficulty.

As for an expected level of contrastive feature use across school ages, our data from Sample 2, the broader, more representative sample, clearly indicate that despite progressive decreases among AAE speaking children at each subsequent age, zero marking and other variable use forms notably did not disappear at the oldest ages. Indeed, being a user of variable use forms was generally more common in our sample than full overt marking and GAE-like usages. Very high levels of zero marking and substitutions (up to 100%) were established as common with 4-, 5-, and 6-year-old typical speakers. This effect was very robust, so variable use of contrastive features at these ages was less available to provide information about whether language difficulty was also present (Seymour et al., 1998). However, convergent with other research, we observed a reversal after age 7 where overt marking and other GAE forms were used more often than variable usages at the oldest ages in our study. For 17% of the AAE children over age 7, total absence of zero marking and other variable AAE usages was found. Thus, the opposite--total *absence* of overt marking--might be considered a-typical past age 7. However, even in a formal academic environment likely to encourage GAE-like forms in children who can code-switch, typical AAE speaking children in our sample at age 9 did not provide responses that would result in their language being identical to GAE children of the same age. Notably, despite the widespread presence of GAE-like usages at the older ages, typically developing children usually demonstrated some zero marking and other variable forms as well. Therefore, variable feature use appeared characteristic of typical development. Further study of children identified as a-typical will be needed to establish whether there are differences sufficiently pronounced to be used as discriminators. (See Jackson & Pearson, 2008, 2010.)

Region and parent education level appeared to influence how quickly dialect density as measured by the DDR decreased, but the effect size for those variables was much smaller than for age. The pattern of small, significant differences from these comprehensive data may explain why some previous studies found differences where others did not. These results suggest that although differences overall were not large, region and SES should be taken into account when possible when examining dialect density, because some specific subgroups may show a local effect. On the other hand, it appears quite justified to collapse across genders.

Limitations

Despite the great volume of data available here, the study had several limitations. These results which were derived from a formal testing situation may have underestimated the use of variable usages for the features examined. Additionally, reliance on EurA examiners may have contributed to a further underestimation of zero marking and predicted substitutions as some participants may have shifted to a more formal register than they typically use (Agerton & Moran, 1995; Pearson et al., 2009; Ratusnik & Koenigsknecht, 1975).

Another limitation arises from the difficulty of making inferences about development over time from cross-sectional data. Longitudinal follow-up studies with this methodology will be required to determine the extent to which the overall pattern of development derived from the patterns for the group are reflected in developmental patterns for individual children.

Finally, all studies are limited in their findings to what they observe. As noted in the literature review, many studies of dialect density are based on spontaneous speech samples (cf. Oetting & McDonald, 2002; Washington & Craig, 1994). These studies almost all indicate that zero copula/auxiliary was one of the most persistent features. In these data based on a very restricted set of items with a small number of elicitation frames, however, the zero copula/auxiliary feature was not among the most persistent. Zero copula/auxiliary may well be part of these children's repertories in spontaneous speech, but it was not among the most frequently *elicited* features.

Directions for Future Research

In order to focus on typical patterns, this study excluded children identified as language impaired. Ultimately, though, one must explore these same patterns among children with language impairment, both AA and EurA. It is well documented that EurA children with language impairment frequently have a prolonged period of omitting certain morphosyntactic forms, sometimes referred to as a period of *Extended Optional Infinitive* (Leonard, 1998; Rice & Wexler, 1996). These forms are frequently identical to the zero marked forms used variably in

AAE. Examining how these zero marked forms manifest in both varieties within their own language impaired groups could yield more explanatory data related to the nature of language impairment in the context of whether the child is acquiring a variety with required overt marking, or a variety that accepts variable marking. (Such work is currently underway in our lab.)

Assessment

Researchers like Oetting and McDonald (2001), who recognize the difference versus disorder conundrum, nevertheless have expressed the opinion that the right combinations of contrastive variable use features have diagnostic potential. The current analysis does not make the crucial (direct) comparisons to respond directly to those claims—namely a comparison of typical and a-typical groups on these same measures. Nonetheless, several features, including features signaled by Oetting & McDonald as potential discriminators of impairment, showed levels of use at or close to 100% among typically developing 4- to 6-year-old children, so there was no room for levels to be higher among an impaired group. On the other hand, exclusive use of zero-marking and AAE substitutions in the older ages may be a red flag that indicates the need for further examination, including administration of a non-contrastive assessment. Unlike in 2001 when Oetting and McDonald found too few candidates for non-contrastive diagnostic indicators, there is currently the possibility of using a full range of non-contrastive elements for such diagnosis (Seymour & Pearson, 2004), and so there is less pressure to find diagnostic potential in contrastive features. Still, it is also possible that a “spectrum” of impairment types exists in which some child speakers’ are more vulnerable to morphosyntax disruptions which would more often affect variable use features, while other child speakers may be more vulnerable to difficulty or impairments localized around the non-contrastive aspects of language (e.g., articles, *wh*-questions, quantifiers) described in the volume edited by Seymour and Pearson

(2004). In this case, finding diagnostic potential in contrastive features would remain a relevant endeavor.

Educational Implications

While the features detailed here are not appropriate targets for *therapeutic* intervention, they may be helpful in situations where typically developing AA children are learning "code-switching" or "style shifting" between home and school language styles. Because of the variable nature of these forms, most AAE-speaking children already use GAE-like alternatives some of the time. For example (as shown in Figure 2), they may be more likely to use 3rd -*s* subject-verb agreement for auxiliary verbs (e.g., *have/has*) than for lexical verbs (e.g, *eat/eats*). Given this, an educator or consulting clinician can use a child's own productions to begin to build awareness of the linguistic contexts in which the child may wish to choose GAE-like forms (i.e., code-switch or style shift). Clinicians can also serve as the educated source for the classroom teacher wanting input on expected trends or appropriate non-biased code-switching techniques.

It has long been a concern that children who do not completely adopt the GAE model in the academic setting may be predisposed to academic difficulty. Despite high levels of AAE variable use morphology and phonology across all ages studied, none of the participants here were identified as high risk. Thus, these data indicate that difficulty that arises from a mismatch between home and school dialects for AAE speakers (Charity et al., 2004) is not due to general incompetence in language, although greater facility in code-switching may be called for in the academic setting.

References

Agerton, E. & Moran, M. (1995). Effects of race and dialect of examiner on language samples elicited from southern African American preschoolers. *Communication Disorders Quarterly, 16,* 25-30.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150-177.

Bountress, N. (1983). Effect of segregated and integrated educational settings upon selected dialectal features. *Perceptual & Motor Skills*, *57*, 71-78.

Charity, A. (2007). Regional differences in low SES African-American children's speech in the school setting. *Language Variation and Change*, *19*, 281-293.

Charity, A., Scarborough, H., & Griffin, D. (2004). Familiarity with school English in African American children and its relation to early reading achievement. *Child Development, 75(5),* 1340-1356.

Coles-White, D'J. R, de Villiers, J. G., & Roeper, T. (2004). The emergence of barriers to *wh*-movement, negative concord and quantification. In A. Brugos, L. Micciulla, & C. Smith (Eds.), *BUCLD 28, Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 98-107). Somerville, MA: Cascadilla Press.

Craig, H. K., Thompson, C., Washington, J. A., & Potter, S. (2003). Phonological features of child African American English. *Journal of Speech, Language, and Hearing Research*, *46*, 623-635.

Craig, H. K. & Washington, J. A. (2002). Oral language expectations for African American preschoolers and kindergartners. *American Journal of Speech-Language Pathology*, *11*, 59-70.

Craig, H. K. & Washington, J. A. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research*, *47*, 450-463.

Craig, H. K. & Washington, J. A. (2006). *Malik goes to school: Examining the language skills of African American students from preschool—5th grade.* Mahwah NJ: Erlbaum.

DeBose, C. E. (1992). Codeswitching: Black English and Standard English in the African-American linguistic repertoire. *Journal of Multilingual and Multicultural Development, 13*, 157-167.

Green, L. J. (2002). *African American English: A linguistic introduction.* Cambridge: Cambridge University Press.

Green, L. J. & Roeper, T. (2007). The acquisition path for aspect: Remote past and habitual in child African American English. *Language Acquisition, 14*, 269-313.

Haynes, W. & Moran, M. (1989). Cross-sectional developmental study of final consonant production in southern Black children from preschool through third grade. *Language, Speech, and Hearing Services in Schools*, *20*, 400-406.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74,* 1368-1378.

Horton-Ikard, R. & Miller, J. F. (2004). It is not just the poor kids: The use of AAE forms by African-American school-aged children from middle SES communities. *Journal of Communication Disorders, 37*, 467-487.

Isaacs, G. J. (1996). Persistence of non-standard dialect in school age children. *Journal of Speech, Language, and Hearing Research*, *39*, 434-441.

Jackson, J. E. (1998). *Linguistic aspect in African-American English speaking children: An investigation of aspectual "be"*. Unpublished Ph.D., University of Massachusetts.

Jackson, J. E. & Green, L. J. (2005). Tense and aspectual *be* in child African American English. In H. Verkuyl, H. DeSwart., & A. Van Hout (Eds.), *Studies in Theoretical Psycholinguistics*: Vol. 32. *Perspectives on Aspect* (pp. 233-250). Dordrecht, Netherlands: Springer.

Jackson, J. E. & Pearson, B. Z. (2008, November). *Non-mainstream dialect and the identification of risk for language impairment.* Paper presented to the annual convention of the Speech, Language, and Hearing Association: Chicago, IL.

Jackson, J. E. & Pearson, B. Z. (2010). Variable use of features associated with African American English by typically-developing children, ages 4 to 12. *Topics in Language Disorders, 30 (2),* 135-144.

Labov, W. (1970). The logic of non-standard English. In F. Williams (Ed.), *Language and poverty* (pp. 153-189). Chicago: Markham Publisher.

Leonard, L. (1998). *Children with Specific Language Impairment.* Cambridge: MIT Press.

Newkirk, B., Stockman, I., Guillory, B. & Seibert, M. (2003, November). *Dialect density and standardized test outcomes: Pilot study.* Paper presented to the annual convention of the Speech, Language, and Hearing Association: Chicago, IL.

Oetting, J. B. & McDonald, J. L. (2001). Non-mainstream dialect use and Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 44,* 207-223.

Oetting, J. B. & McDonald, J. L. (2002). Methods for characterizing participants' nonmainstream dialect use in child language research. *Journal of Speech, Language, and Hearing Research, 45,* 505-518.

Ogbu, J. (2003). *Black American students in an affluent suburb: A study of academic disengagement*. Mahwah, NJ: LEA.

Pearson, B. Z., de Villiers, P., Magaziner, K., & Sunderland , K. (2005, November). *Validation of the DELV-NR by Language Sample Analysis.* Poster presented at the American Speech Language and Hearing Association Annual Meeting, San Diego, CA.

Pearson, B. Z., Velleman, S. L., Bryant, T. J. & Charko, T. (2009). Phonological milestones for African American English-speaking children learning Mainstream American English as a second dialect. *Language, Speech, and Hearing Services in Schools, 40 (3),* 229-244.

Pollock, K. & Berni, M. C. (1996). *Vocalic and postvocalic /r/ African American Memphians*. Paper presented at the meeting of New Ways of Analyzing Variation in English (NWAVE), Las Vegas, NV

Ramer, A. & Rees, N. (1973). Selected aspects of the development of English morphology in Black American children of low socioeconomic backgrounds. *Journal of Speech and Hearing Research*, *16*, 569-577.

Ratusnik, D. & Koenigsknecht, R. (1975). Influence of certain clinical variables on Black preschoolers' non-standard phonological and grammatical performance. *Journal of Communication Disorders*, *8*, 281-297.

Renn, J. & Terry, J. M. (2009). Operationalizing style: Quantifying the use of style shift in the speech of African American adolescents. *American Speech, 84,* 367-390.

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.

Seymour, H., Bland-Stewart, L. & Green, L. (1998). Difference versus deficit in child African-American English. *Language, Speech, and Hearing Services in Schools, 29*, 96-108.

Seymour, H. N. & Pearson, B. Z. (Eds.) (2004). *Seminars in Speech and Language, 25 (1).*

Seymour, H. N., Roeper, T., & de Villiers, J. G. (2000). *Dialect Sensitive Language Test (DSLT), (unpublished pilot edition).* San Antonio TX: The Psychological Corporation.

Seymour, H. N., Roeper, T., & de Villiers, J. G. (2003). *Diagnostic Evaluation of Language Variation, Screening Test (DELV-ST).* San Antonio TX: The Psychological Corporation.

Seymour, H. N., Roeper, T., & de Villiers, J. G. (2005). *Diagnostic Evaluation of Language Variation, Norm Referenced (DELV-NR).* San Antonio TX: The Psychological Corp.

Stockman, I. J. (1996). The promise 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. (2010). A review of developmental and applied language research on African American children: From a deficit to difference perspective on dialect differences. *Language, Speech, and Hearing Services in Schools*, *41* (1), 23-38.

Tager-Flusberg, H., & Cooper, J. (1999). Defining the phenotype of SLI. *Journal of Speech, Language, and Hearing Research, 42*, 1275–1278.

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

U.S. Bureau of the Census. (2000). *Current Population Survey, October 2000: School Enrollment Supplemental File* [CD-Rom]. Washington DC: Bureau of the Census.

Velleman, S. J. & Pearson, B. Z. (2010). Differentiating speech sound disorder from phonological dialect difference: Implications for assessment and intervention. *Topics in Language Disorders, 30 (3).*

Von Hofwegen, J. & Wolfram, W. (in press). Coming of age in African American English: A longitudinal study. *Journal of Sociolinguistics.*

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

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

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

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

Wolfram, W. A. (1991). *Dialects in American English*. Englewood Cliffs, NJ: Prentice Hall.

Appendix

Contrastive Variable Use Item Types and Tokens for Steps 1 and 2

ITEM TYPES	Examples	TOKENS Step 1	Step 2 "core"
MORPHOSYNTAX			
Variable Zero Marking			
Zero present tense copula			
is copula	*He ∅ tall.* *He ∅ a doctor.*	*2*	--
Zero present tense auxiliary			
is auxiliary	*He ∅ running.*	2	--
are auxiliary	*They ∅ sleeping.*	2	--
Zero plural /–*s*/	*two glass∅*	2	--
Zero possessive /–*s*/	*John'∅ mother left.*	3	--
Zero regular past tense /–*ed*/	*He play∅ yesterday.*	3	--
Variable Agreement & Negation			
Subject verb 3rd –*s* with *do*	*He don't like to swim.*	*5*	*2*
Subject verb 3rd –*s* with *have*	*She have no shoes.*	*2*	*2*
Subject verb 3rd –*s* with lexical verb	*He sleep∅ on a bed.*	*5*	*4*
Subject verb past copula/ auxiliary	*They was cold.* *They was sleeping.*	*5*	*2*
Multiple negation	*He don't have no shoes.*	3	--
PHONOLOGY			
Final consonant clusters	*test* pronounced [t□s]	*15*	*1*
Voiced interdental fricatives /δ/ in any position	*this* pronounced [dIs]; *breathe* pronounced [briv]	*5*	*2*
Voiceless interdental fricatives /θ/ in final position	*moth* pronounced [maf]	3	2
Other final consonants	*asleep*	*32*	--

(from Seymour & Pearson, 2004)

Table 1

Sample 1: (All Typically-developing)

	Age in Years	4;0 – 4;11	5;0 – 5:11	6;0 – 6:11	7;0 – 8:11	9;0 – 10:11-	11;0 – 12:11	Total
AAE	*N*	80	99	118	39	62	45	**443**
	Parent Education	1.9	2.1	2.0	2.1	2.1	2.1	2.0
	% Female							55%
	Region:	North Central		South		North East		West
	From Census, AA population	25%		51%		15%		9%
	Sample 1	30%		63%		6%		3%
	Parent Education Level	0-11 yrs		(h.s. degree)		13-15		(college degree)
	From Census, AA population	18%		38%		34%		10%
	Sample 1	18%		60%		19%		3%
GAE	*N*	42	42	51	19	15	25	**194**
	Parent Education	2.2	2.1	1.9	2.1	2.4	2.5	2.2
	% Female							55%
	Region:	North Central		South		North East		West
		20%		53%		9%		10%
ALL	N	122	141	169	58	77	70	**637**

AAE: Speaker of African American English as a first dialect, AA ethnicity

GAE: Speaker of General American English, European-American background (EurA)

Average Parent Education Level (PED): On a 4-point scale adapted from The Psychological Corporation, 1 represents 0 to 11 years, no high school degree, 2 a high-school degree, and 4 a college degree.

Table 2

Sample 2, Preliminary African American NIH Norming Sample (All Typically-Developing)

Age in Years	4;0 – 4:11	5;0– 5:11	6;0 – 6:11	7;0 – 7:11	8;0 – 8;11	9;0 – 9;11	Total
N =	150	143	158	83	70	83	**687**
Parent Ed.	2.4	2.5	2.5	2.4	2.4	2.4	2.45
% Female							51.5%
Region	North Central		South		North east		West
Region from AA Census	25%		51%		15%		9%
Region, Sample 2	23%		50%		18%		9%
	0-11 yrs		h.s. degree		13-15 yrs		16+ years
PED Level, AA Census	18%		38%		34%		10%
PED Level, Sample 2	15%		37%		35%		13%

Average Parent Education Level (PED): On a 4-point scale adapted from The Psychological Corporation, 1 represents 0 to 11 years, no high school degree, 2 a high-school degree, and 4 a college degree.

Table 3

Combined African American Sample (Sample 1 ages 4-9 and Sample 2)

Age in Years	4;0 – 4:11	5;0– 5:11	6;0 – 6:11	7;0 – 7:11	8;0 – 8;11	9;0 – 9;11	Total
N =	234	243	287	99	95	112	**1051**
REGION by Parent Education by Age							
Parent Ed.	North Central						275 (26%)
0-11 years	17	10	16	5	4	4	
h.s. degree	28	40	30	5	14	12	
13-15 years	15	18	18	7	7	7	
4-year degree	4	7	2	2	2	1	
	South						557 (53%)
0-11 years	20	16	29	11	7	9	
h.s. degree	56	57	76	17	26	31	
13-15 years	32	35	37	19	10	15	
4-year degree	12	11	14	6	6	5	
	North-east						146 (14%)
0-11 years	8	4	2	3	2	4	
h.s. degree	8	14	15	7	3	6	
13-15 years	11	10	14	5	5	6	
4-year degree	3	5	5	1	2	3	
	West						73 (7%)
0-11 years	0	1	0	1	1	0	
h.s. degree	7	7	6	3	2	2	
13-15 years	8	7	10	3	3	5	
4-year degree	1	0	2	2	1	1	

Parent Education Level (PED): (Adapted from The Psychological Corporation)

Table 4

Mean Dialect-Density Ratio (DDR) by Age, Region, Parent Education (PED) Level, and Gender (Sample 2)

Age	Mean (SD)	Region	Mean (SD)	PED Level	Mean (SD)	Gender	Mean (SD)
4 yrs	.728 (.27)	**North Central**	.629 (.34)	**0-11 years**	.604 (.34) [c]	**Male**	.522 (.34) [d]
5	.624 (.28)	**South**	.528 (.33)	**h.s. degree**	.571 (.33) [c]	**Female**	.510 (.34) [d]
6	.570 (.32)	**Northeast**	.428 (.32)	**13-15 years**	.400 (.32)		
7	.386 (.28) [a]	**West**	.319 (.32)	**4-yr degree**	.299 (.29)		
8	.347 (.29) [a b]						
9	.278 (.28) [b]						

Values with same-letter superscript are within the 95% confidence interval of each other.

Alternate Table 2

Table 2

Sample 2, Preliminary African American NIH Norming Sample (All Typically-Developing) and Samples 1 and 2 Combined

Age in Years	4;0 – 4:11	5;0– 5:11	6;0 – 6:11	7;0 – 7:11	8;0 – 8;11	9;0 – 9;11	Total
Sample 2 *N* =	150	143	158	83	70	83	**687**
Samples 1 + 2 *N* =	234	243	287	99	95	112	**1051**
% Female							
Sample 2							51.5%
Samples 1 + 2							53%
Region	North Central		South		North east		West
Region from AA Census	25%		51%		15%		9%
Region, Sample 2	23%		50%		18%		9%
Samples 1 + 2	26%		53%		14%		7%
	0-11 years		h.s. degree		13-15 years		16+ years
PED Level, AA Census	18%		38%		34%		10%
PED Level, Sample 2	15%		37%		35%		13%
PED Level, Samples 1+ 2	17%		45%		28%		9%

Figure Captions

Figure 1. *Percent of Variable AAE Usage of Contrastive Morphosyntax and Phonology Features (of 89) by Age and Dialect Group (Sample 1)*

Figure 2. *Percent of Variable AAE Usage for Four Most Contrastive Morphosyntax Features by Age (GAE control group by overall average, all ages)*

Figure 3. *Percent of Variable AAE Usage for Four Less Contrastive Morphosyntax Features by Age (GAE-control group by overall average, all ages)*

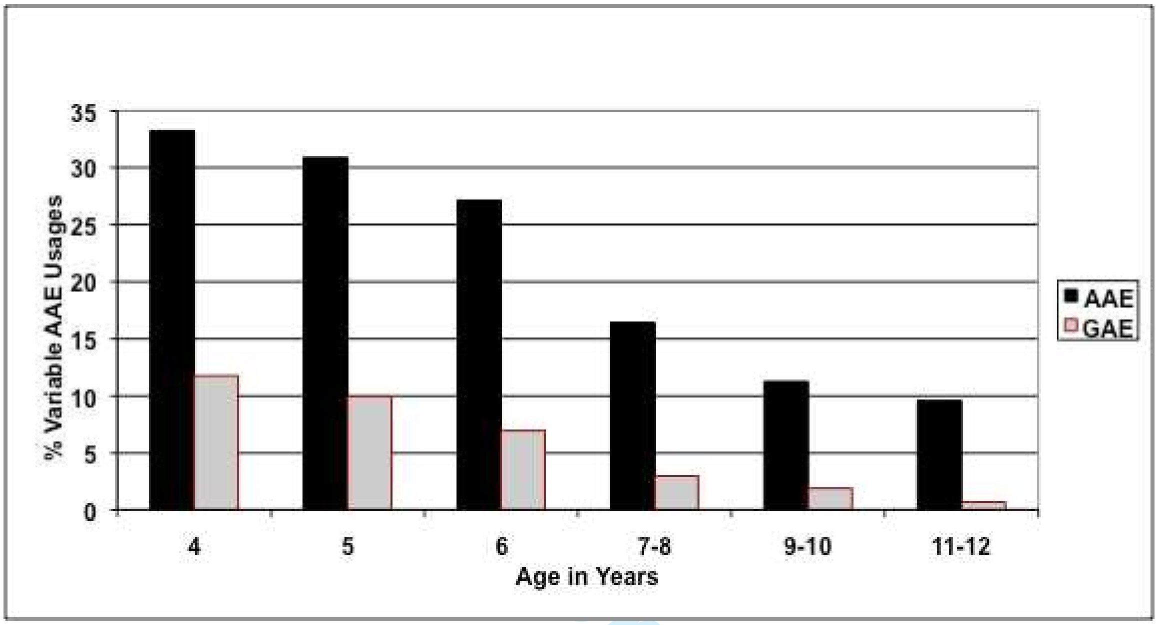
Figure 4. *Percent AAE Usage for Three Most Contrastive Phonological Features by Age (GAE control group by overall average, all ages)*

Figure 5. *Dialect Density Ratio (DDR) Mean and Standard Deviations by Age (Sample 2)*

Figure 6. *Dialect Density Ratio (DDR) by Age and Parent Education Level (Combined Sample)*

Figure 1

Percent of Variable AAE Usage of Contrastive Morphosyntax and Phonology Features (of 89) by Age and Dialect Group (Sample 1)

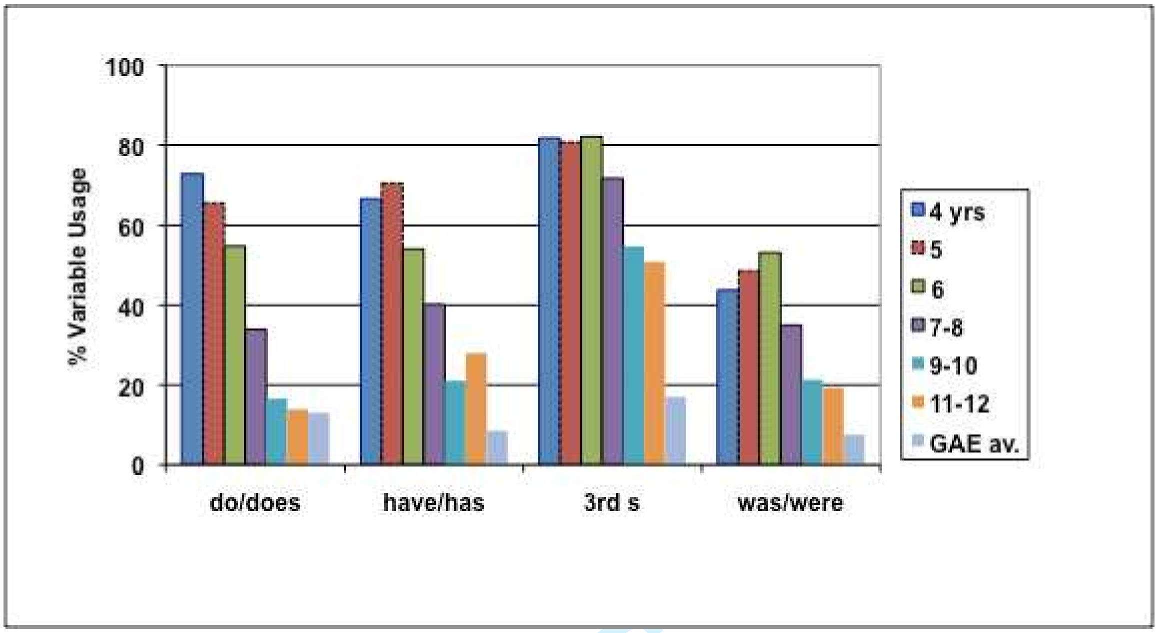


AAE = African American English
GAE = General American English

Figure 2

Percent of Variable AAE Usage for Four Most Contrastive Morphosyntax Features by Age

(GAE control group by overall average, all ages)

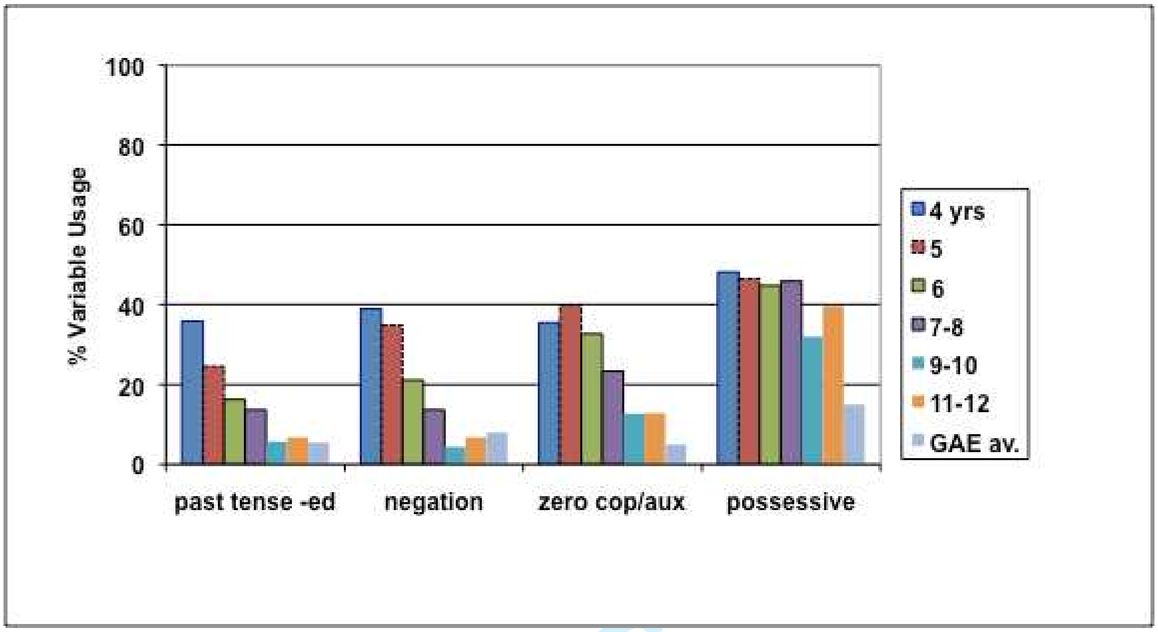


AAE = African American English
GAE = General American English

Figure 3

Percent of Variable AAE Usage for Four Less Contrastive Morphosyntax Features by Age

(GAE control group by overall average, all ages)

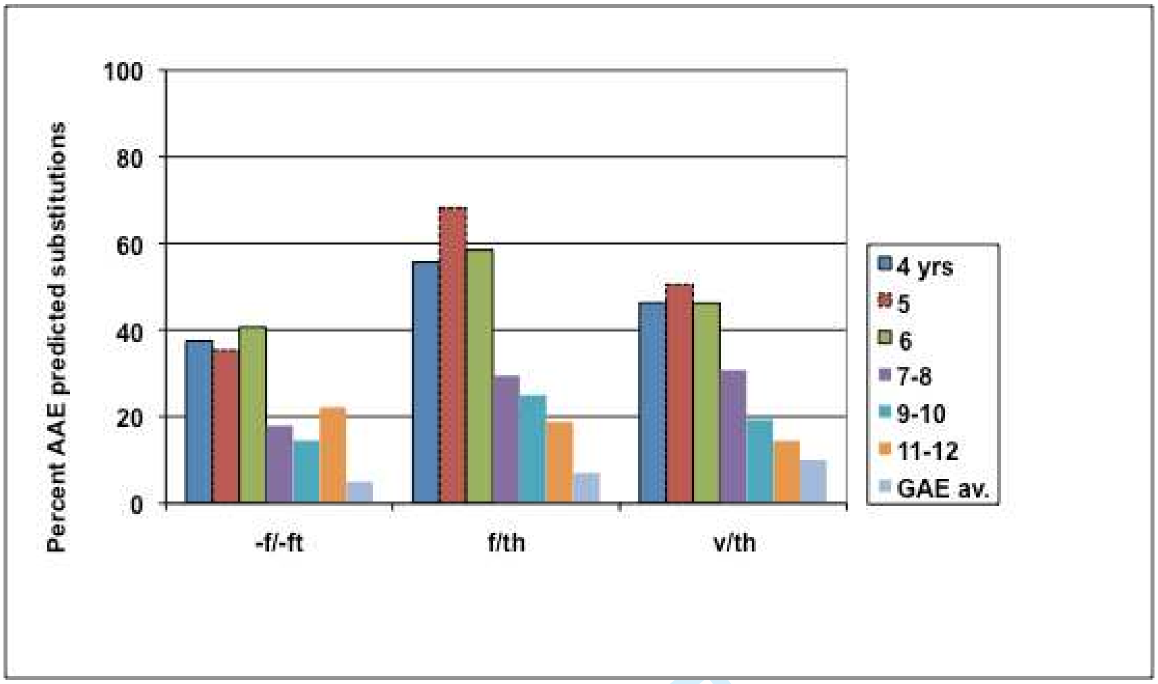


AAE = African American English
GAE = General American English

Figure 4

Percent of Variable AAE Usage for Three Most Contrastive Phonological Features by Age

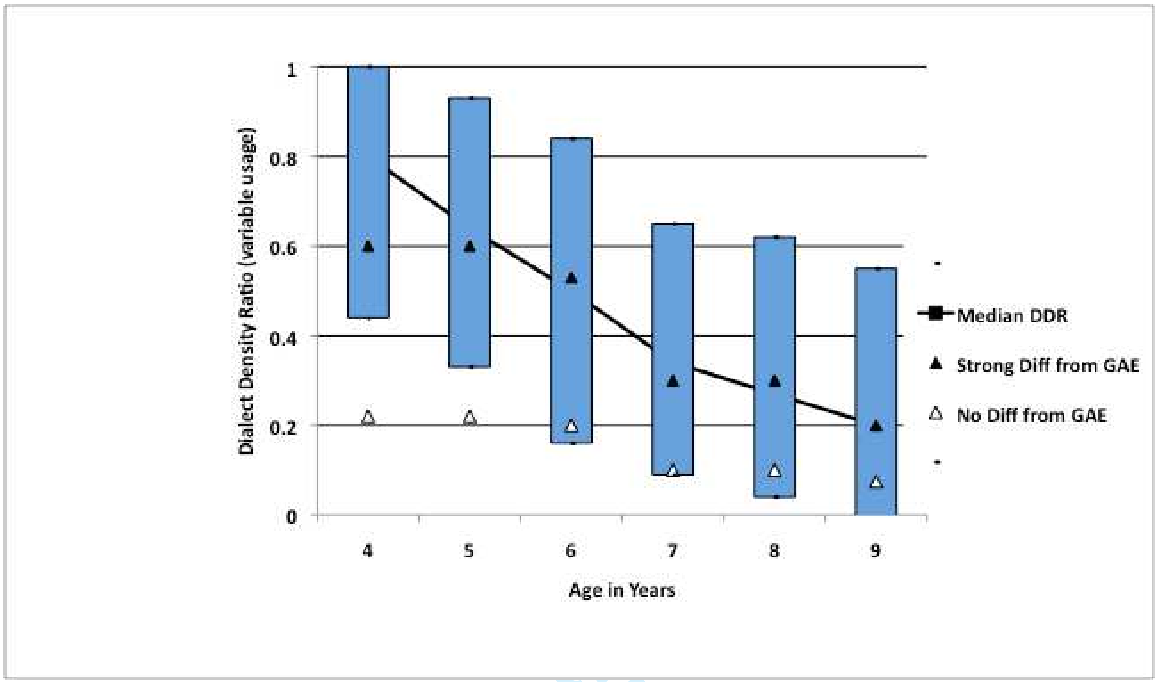
(GAE control group by overall average, all ages)



AAE = African American English
GAE = General American English

Figure 5

Dialect Density Ratio (DDR) Mean and Standard Deviations by Age (Sample 2 alone)



Bars indicate plus and minus one standard deviation from the mean

"No" and "Strong Difference from General American English (GAE)" benchmarks from European American (EurA) participants in Sample 1

Figure 6

Dialect Density Ratio (DDR) by Age and Parent Education Level (Combined Sample)