Published in final edited form as:
J Speech Lang Hear Res. 2013 June ; 56(3): 933–944. doi:10.1044/1092-4388(2012/12-0099).

Linguistic Constraints on Children's Overt Marking of BE by Dialect and Age

Joseph Roy[1], **Janna B. Oetting**[2], and **Christy Wynn Moland**[2]
[1]University of Ottawa

[2]Louisiana State University

Abstract

Purpose—Overt marking of BE in nonmainstream adult dialects of English is influenced by a number of linguistic constraints, including the structure's person, number, tense, contractibility, and grammatical function. In the current study, we examined the effects of these constraints on overt marking of BE in children as a function of their nonmainstream English dialect and age.

Methods—The data were language samples from 62 children, aged four to six years; 24 spoke African American English (AAE) and 38 spoke Southern White English (SWE). Analyses included analysis of variance and logistic regression.

Results—Rates of overt marking varied by the children's dialect but not their age. Although the person, number, tense, and grammatical function of BE influenced the children's rates of marking, the nature and magnitude of the influence differed by the children's dialect. For AAE-speaking children, contractibility also influenced their marking of BE.

Conclusions—Consistent with the adult literature, the AAE- and SWE-speaking children marked BE in ways that differed from each other and from what has been documented for child speakers of Mainstream American English. These findings show stability in the use of BE in AAE and SWE that spans different generations and different dialect communities.

Adult African American English (AAE) and Southern White English (SWE) differ from Mainstream American English (MAE) in the production of copular and auxiliary forms of BE. Unlike MAE, overt marking of BE is not obligatory for adults who speak AAE or SWE. In fact, it is perfectly acceptable in these two dialects for both overtly marked (e.g., *he is happy*) and zero marked BE (*he Ø happy*) to be produced. However, overt versus zero marking of BE does not occur randomly. Instead, the type of marking produced is probabilistic and tied to the linguistic characteristics of the context surrounding the BE form.

Effects of linguistic contexts on AAE and SWE speakers' overt marking of BE are commonly referred to as linguistic constraints because various contexts are thought to constrain (or influence) when a speaker produces or doesn't produce an overtly marked form (for review of constraint literature and earlier work by Labov and others, see Sankoff, Taliamonte, & Smith, 2005). Linguistic constraints have been documented in a number of studies involving adult speakers of AAE and to a much lesser extent SWE (Bailey & Maynor, 1985; Baugh, 1980; Blake, 1997; Childs & Malinson, 2004; Labov, 1969; 1972; Romaine, 1982; Rickford, Ball, Blake, Jackson, & Martin, 1991; Rickford, 1998; Walker, 2000; Winford, 1992; Wolfram, 1969; 1974). The presence of some linguistic constraints on children's overt marking of BE has also been documented in a few studies. However, these

Contact Author Joseph Roy Department of Linguistics, Room 401 70 Laurier Avenue East University of Ottawa Ottawa, Ontario K1N 6N5, Canada jroy042@uottawa.ca.

studies have often focused on either the copular or auxiliary BE form, and the analyses have either been highly descriptive or limited to analyses of variance. Given this, the goal of the current study was to extend the study of children's overt marking of BE by including copular and auxiliary forms, speakers of two nonmainstream dialects of English (AAE and SWE) and two age groups (four and six years), and by utilizing two types of statistical analyses, analysis of variance and logistic regression.

Although previous adult studies have examined a variety of linguistic constraints, the current work focuses on three: the person, number, and tense of the BE form, the contractibility of the BE form, and the grammatical function of the BE form.[1] These constraints were chosen to maximize the number of BE tokens available for the analyses. Each of these linguistic constraints has also been studied by others with children who speak either AAE or MAE; however, they have not been examined together in a single study nor have they been used to compare different child dialects to each other. Studies of BE marking (and grammar in general) have also not been conducted on children who speak SWE, even though this dialect is spoken by many children who live in the rural South (Wolfram & Schilling-Estes, 1998), and the socioeconomic levels and the academic achievement of children who live in the rural South are lower than those of children who live in other regions of the United States (Harris & Zimmerman, 2003; Miller & Weber, 2004).

A detailed study of BE and the constraints that influence its use in child AAE and SWE is important because zero marking of this structure has been repeatedly shown to be more frequent than any other nonmainstream grammar structure within these dialects (for child AAE and SWE, see Oetting & Pruitt, 2005; for child AAE, see Horton-Ikard & Weismer, 2005; Jackson & Roberts, 2001; Washington & Craig, 1994). Speech-language clinicians who work with diverse groups of children in the rural South (and perhaps elsewhere) have likely heard overtly marked and zero marked forms of BE within their community, yet the field lacks empirical data from children to help clinicians understand and articulate when and why a child might produce an overtly marked form over a zero marked form and vice versa.

As background for the study, we operationally define the three linguistic constraints and review what is known about their effects on adults' marking of BE. Findings from the adult studies provide a baseline from which to consider the children's dialects. If the children's marking of BE is found to be similar to that of previous adult studies, we will have evidence of stability in BE across different generations of speakers and time. If differences are identified, we will have evidence of linguistic evolution and change.

Findings by Wolfram and Thomas (2002) indicate that modern day changes in language are likely to be structure specific and dialect specific. Their study focused on six grammar structures (i.e., *weren't* leveling, *was* leveling, verbal –s marking with third person plural subjects, zero *is*, zero *are*, and zero verbal –s) in the conversations of 49 adults (35 African American and 14 white) who resided in Hyde county, North Carolina. The African American adults reflected four age groups (14–23, 32–43, 55–70, and 77–102 years), and the white adults reflected two (15–27 and 77–94 years). Results showed that the African American and white groups presented varying degrees of change in the six grammar structures. For the African Americans, *weren't* leveling and verbal –s marking showed significant decreases in use across generations, zero *is* showed a significant increase, and

[1]Other linguistic constraints that have been examined in some adult AAE studies include: the preceding subject context (it/that/what vs. other pronouns vs. specific nouns), stress (emphatic vs. non-emphatic), the clause position (final vs. non-final), and the preceding and following phonological context (vowel vs. consonant). For discussion and analysis of various combinations of these constraints in adult AAE, see Blake (1997), Rickford et al. (1991), and Wyatt (1996).

was leveling and zero *are* showed minimal change. In comparison, the white participants showed a significant increase of use across generations in *weren't* leveling and a significant decrease in verbal –s marking with third person plural noun phrases; all other grammar structures showed minimal change. Dubois and Horvath (1998; 1999) have documented similar structure specific changes across generations in phonological patterns that have been historically tied to Cajun English in Louisiana. In their studies, some structures show a decrease in use from older to younger speakers, others show an increase, and still others show a v-shape pattern of change, with high rates of use in older and younger speakers, and low rates of use in middle-age speakers.

In our own work with the same children studied here, we have identified some structure specific and dialect specific changes through separate analyses of children's nonmainstream productions of verbal –s and relative clause markers (Cleveland, 2009; Oetting & Newkirk, 2011). For verbal –s marking, the data show stability across generations for both AAE and SWE when the children's data are compared to previous studies from adults. For relative clause markers, the data show a significant reduction of nonmainstream productions in child AAE when compared to previous adult AAE studies, with stable rates of nonmainstream productions in SWE. The current study allows for further examination of language evolution and change by focusing on children's marking of BE.

Following a review of the adult studies, we present findings from the child language literature. These studies are important for two reasons. First, although these studies have varied in the age of the child examined and in their methodologies, they suggest dialect differences between AAE, SWE, and MAE. In the current work, we directly compare child AAE to SWE while also considering MAE child studies that have employed methods similar to ours. Second, these previous studies highlight the developmental nature of child language and motivate the need to consider age effects within our data. Child studies that have examined age effects repeatedly show higher rates of overt marking of BE in five- and six-year-olds than in three- and four-year-olds, and in MAE, adult levels of BE marking are not expected until children are beyond the age of four years (Rice, Wexler, & Hershberger, 1998). In the current work, we examine the developmental nature of BE marking in AAE and SWE by examining children who are aged four and six years.

Within the literature review, we report all findings from previous studies as a function of the speakers' rates of overtly marked forms as opposed to their rates of zero marked forms. Although the field of sociolinguistics focuses on the zero marked forms, the field of child language focuses on the overtly marked forms. For those interested in rates of zero marking, percentages reported in this paper can be converted using the formula, 1 – the rate of overtly marked forms (e.g., 1 − 40% = 60%).

Another methodological difference across fields is in the way rates of marking are calculated. Within the adult sociolinguistics literature, rates of marking are often calculated by pooling data across speakers (e.g., Blake, 1997; Labov, 1969; Rickford, 1999; Rickford et al., 1991). Thus, rates reflect the proportion of time a particular form is produced out of all possible contexts available, and speakers who produce more data carry more weight in the calculation than those who produce less. In contrast, within the child language literature, rates are typically calculated for each child and then these individual rates are averaged to reflect a mean percentage (e.g., Cleave & Rice, 1997; Garrity & Oetting, 2010; Leonard et al., 2003; Rice et al., 1998). Thus, each child's data contributes to the mean in the same way regardless of the number of tokens produced. In the current work, we specify the method of calculation for each study reviewed as well as consider both methods within the analyses.

Linguistic Constraint: Person, Number, Tense

The first linguistic constraint involved the person, number, and tense of the BE form. As shown in 1–4 below, this variable included four levels based on the grammatical subject and temporal reference of the verb. For the study here, we combined BE tokens for *was* and *were* because there were low numbers of these forms in the data and predictions about rates of overt marking for these forms were the same. Inspection of the data also suggested similar findings for *was* and *were*.

1 Am: First person, singular, (temporal referent: present/habitual)

2 Is: Third person, singular, (temporal referent: present/habitual)

3 Are: Second person and first and third person plural, (temporal referent: present/habitual)

4 Was/were: (temporal referent: past, past counterfactual)

The adult dialect literature repeatedly shows that both AAE and SWE speakers are most likely to show variable marking (i.e., both overt marking and zero marking) of BE forms that involve *are* but rarely show variable marking of BE when the forms involve *am*, *was*, or *were* (Blake, 1997; Labov, 1969; Rickford, 1999; Rickford et al., 1991; Wolfram, 1974). For these latter three contexts, both AAE and SWE adult speakers produce such high rates of overt marking that Blake (1997) and others view these contexts as categorically non-variable and recommend that they be excluded from analyses.

The adult literature shows that AAE speakers also variably mark *is* in ways that are similar to *are*; however, rates of overt marking are typically higher for *is* than *are*. For example, in Rickford et al.'s (1991) study of *is* and *are* within AAE, *is* was overtly marked 47% of the time compared to *are*, which was overtly marked 22% of the time. These calculations pooled data from 30 speakers and were proportional in nature.

For adults who speak SWE, the literature is less clear. Labov (1969) and Fasold and Wolfram (1970) claim that white southerners do not zero mark *is* in the variable manner that has been documented for varieties of AAE, but Wolfram (1974), Cukor-Avila (2001), and Hazen (2001) present data from southern white speakers which show some variable marking of *is*. In all cases, however, SWE-speaking adults have been shown to overtly mark *is* (and even *are*) at relatively high rates. For example, rates of overt marking by Hazen's (2001) white speakers from North Carolina were 96% for *is* and 84% for *are*. As evidenced by these data, the magnitude of the constraint effect in adult SWE appears minimal, but the direction of the effect (*is* > *are*) is consistent with the adult AAE literature.

Linguistic Constraint: Contractibility

The second constraint was the contractibility of the BE context. To be consistent with other child language studies (e.g., Cleave & Rice, 1997; de Villiers & de Villiers, 1973; James & Kahn, 1982; Johnston & Schery, 1976), we followed Brown (1973) and coded contexts as contractible if contractibility was possible, regardless of whether the child actually contracted the BE form. Brown's rationale for this decision was based on the difficulty of coding contractibility in a reliable manner across examiners when the data are samples of children's conversational speech. Thus, utterances such as, *They are walking* and *They're walking* were coded as contractible, and utterances such as *Sis is walking* and *Yes she is* were coded as uncontractible for phonetic and/or syntactic reasons. In the former example, *Sis* and *is* cannot be contracted and produced as a single syllable for phonetic reasons; in the latter example, *she* and *is* cannot be contracted and produced as a single syllable for syntactic reasons. When children zero marked a BE form, the contractibility of the context

was determined based on the contractibility of the structure had it been overtly marked. Thus, *Now they Ø walking* was coded as contractible and *This Ø being put right here* was coded as uncontractible.

The adult dialect literature is inconsistent in the coding of contractibility and zero marking of BE. In fact, since Labov's (1969) seminal study of BE, sociolinguists have debated not only the processes and order by which speakers contract and zero mark various forms of BE but also the formulas used to calculate rates of contraction and zero marking within and across various speaker groups (Blake, 1997; Rickford et al., 1991). As an example, Rickford et al. (1991) identifies three different calculations for determining the rates at which a speaker contracts BE and two different calculations for determining when a speaker zero marks BE. Although the effect of contractibility on BE marking (overt vs. zero) has not been the focus of the adult studies, Labov, Rickford, and others have shown that for AAE, zero marking is most likely to occur in contexts that are contracted. Given this, overt marking in adult AAE is more likely to occur in uncontractible contexts as compared to contractible contexts.

In most studies of SWE, the effect of contractibility on speakers' overt marking of BE has not been examined, and presumably this is because rates of overt marking across contractible and uncontractible contexts are high and invariable. Data by Hazen (2001) is consistent with this claim because in his study, SWE speakers did not vary their overt marking of BE by the contractibility of the form.

Linguistic Constraint: Grammatical Function

The third constraint was the grammatical function of the BE structure, and this variable involved two levels to represent copular functions (e.g., *T'Lane is a girl*) and auxiliary functions (e.g., *T'Lane is walking*). Within the adult literature, grammatical function has not been studied in this way. Instead, rates of overt marking have been examined in five different contexts which vary in the content that follows the BE form. As shown in 5–9 below, three of these contexts reflect a copular BE function and two reflect an auxiliary function.

5 *The book is on the table*: Locative context following BE (copular)

6 *The book is brown*: Adjective context following BE (copular)

7 *Her name is T'Lane*; Noun context following BE (copular)

8 *T'Lane is gonna read it*: *Gonna* following BE (auxiliary)

9 *T'Lane is going to the store*: Verb + ing following BE (auxiliary)

In the adult AAE literature, rates of overt marking have been documented to be higher in copular contexts than in auxiliary contexts. Again using AAE data from Rickford et al. (1991) as an example, factor weights which were derived from proportional data were above .50 for all three copular contexts and below .40 for the two auxiliary contexts. Similarly, factor weights provided by Wolfram and Thomas (2002) for African Americans were higher for the three copular contexts (> .46) than for the two auxiliary contexts (.24 and .35). Across both of these AAE adult studies, the factor weights indicate higher rates of overt marking in copular contexts as compared to auxiliary contexts. Limited data exist for adult speakers of SWE; however, Hazen (2001) reports that his SWE speakers overtly marked BE in copular contexts at higher rates than auxiliary contexts (96% vs. 78%). Again, the direction (copular > auxiliary) of the effect in adult SWE is consistent with the adult AAE literature.

Children's Marking of BE

Children's marking of BE has also been documented in some studies, and at least two have included AAE-speaking children. Wyatt (1991) studied overt marking of copular BE in 10 AAE-speaking children, aged 3 to 5 years. Consistent with the adult literature, the children overtly marked *is* and *are* at variable rates, and their rate of overt marking for *is* was higher than their rate for *are* (81% vs. 55%). Similar to studies completed in the adult literature, rates of use were calculated as the proportion of overtly marked BE contexts.

Garrity and Oetting (2010) examined overt marking of auxiliary *am*, *is*, and *are* in 30 AAE-speaking children, aged 4 to 6 years, and 20 of these children were classified as typically developing.[2] The data came from language samples and an elicitation probe that included 30 items (10 for each BE structure). In contrast to the adult studies and Wyatt's study, rates of overt marking were calculated for each child and then averaged.

Using data from children who produced at least three BE contexts within their language samples, rates of overt marking were higher for *am* (4-year-olds = 96%; 6-year-olds = 100%) as compared to *is* (4-year-olds = 26%; 6-year-olds = 55%) and *are* (4-year-olds = 1%; 6-year-olds = 39%). For the elicitation probe, however, rates of overt marking for all forms of BE were high, ranging from 70–90%, and these rates of overt marking were not found to differ by the person, number, or tense of the BE form. There was also a statistically significant age difference in rates of marking for *are* and a marginally significant difference for *is* (4-year-olds < 6-year-olds) when the language samples but not elicitation probe data were examined.

Finally, Leonard and colleagues examined the effects of person, number, and tense on children's rates of overt marking in two studies that included speakers of MAE (Leonard et al., 2003; Polite & Leonard, 2007). The first included 45 children, aged 3 to 5 years, and the second included 32 children, aged 4 to 6 years. Data were limited to auxiliary BE forms which were elicited using probes with 18 items for *am* and 16 items for *is* and *are* combined and *was* and *were* combined. Similar to Garrity and Oetting (2010), the children's individual rates of marking were calculated and then averaged.

In Polite and Leonard (2007), rates of overt marking for *am* were above 90% for both age groups of typically developing children. In Leonard et al. (2003), an age but not BE form effect was observed when the children's combined rates of overtly marked *is* and *are* were compared to their combined rates of overtly marked *was* and *were*. The 5-year-olds' combined rates of *is* and *are* marking and *was* and *were* marking were 84% and 97%, respectively; and the 3-year-olds' combined rates were 81% and 78%, respectively. These findings show developmental changes in rates of marking between 3 and 5 years which is consistent with age effects identified in Garrity and Oetting (2010). Nevertheless, these findings also show that typically developing MAE-speaking children overtly mark auxiliary BE at high rates regardless of the person, number, and tense of the BE form.

Although the effects of contractibility and grammatical function were not the focus of the above studies, these variables have been examined in other studies that have included MAE-speaking children (e.g., de Villiers & de Villiers, 1973; James & Kahn, 1982; Johnston & Schery, 1976). As a recent example, Cleave and Rice (1997) examined BE marking in 10 MAE-speaking three-year-olds. The data were from language samples which were similar to those collected in the current study. Collapsed for grammatical function, the children's

[2]This study as well as a few other child language studies reviewed includes children classified as specifically language impaired; findings from these children are not discussed because they are not the focus of the work.

average rates of overt marking were higher for contractible contexts than for uncontractible contexts, although rates of overt marking for both contexts were relatively high (>75%). Collapsed for contractibility, the children's average rates of overt marking were higher for copular contexts than for auxiliary contexts, although again rates of overt marking for both contexts were relatively high (>80%).

Although linguistic constraints have been documented in young MAE speakers' marking of BE, by the age of six years, effects for contractibility and grammatical function on MAE-speaking children's marking of BE likely disappear. Data to support this claim come from longitudinal work by Rice, Wexler, and Hersberger (1998). In their study, 41 typically developing MAE-speaking children, aged 2;6 to 8;9, were administered language samples and probes at six-month intervals. By the age of six, the children's average percentage of overt marking for auxiliary BE was at or near 100%. This high rate of overt marking in BE contexts by MAE-speaking six-year-olds would be difficult to achieve if the appearance of these overtly marked forms were linguistically constrained by contractibility or grammatical function.

In summary, the linguistic constraints of person, number, tense, contractibility, and grammatical function have been examined in a number of adult and child studies, and these studies have included various groups of AAE, SWE, and MAE speakers. Although the age of the participants and the focus, data, and methods of these studies have varied, findings suggest dialect differences for the constraints reviewed. For AAE, the adult literature suggests that all three constraints should affect the children's rates of overt marking, with the direction of the effects showing: *am* and *was/were* > *is* > *are*; uncontractible > contractible; copular > auxiliary. For children who speak SWE, the adult literature suggests effects for the constraint of person, number, and tense, with the direction of the effect showing: *am, was/were* and *is* > *are* (with possibly a slightly lower rate of marking for *is* than *am* and *was/were*). The literature also suggests that the effects of contractibility on child SWE should be absent, and effects of grammatical function, if present, should parallel the effects observed for child AAE by showing copular > auxiliary. If these predictions are confirmed, the findings will demonstrate stability of BE marking across different generations of AAE and SWE speakers and dialect differences between child AAE and SWE. These findings will also demonstrate dialect differences between these two child dialects and child MAE given the relatively high rates at which child MAE speakers overtly mark all forms and contexts of BE. Note also that when effects for contractibility have been documented for child MAE, the direction of the effect (contractible > uncontractible) has differed from what has been documented in the adult nonmainstream English literature (uncontractible > contractible).

Method

Data

The data were spontaneous language samples from 62 children who had previously been studied by Oetting and McDonald (2001, 2002; see also Oetting & Garrity, 2006; Oetting & Newkirk, 2008; 2011; Oetting & Pruitt, 2005). All of the children lived in a rural area of Southeastern Louisiana and attended either public kindergartens or child development centers and Head Starts in close proximity to the kindergartens. For the 43 families who provided parental occupation and/or educational levels, estimates of social strata were based on Hollingshead (1975). Parental occupations ranged from skilled craft, clerical, and sales groups to small business, minor professional, and technical groups; maternal education level averaged 13 years (SD = 1.80; median = 12 years; range = 9 – 16 years).

Twenty-four children were classified as African American and speakers of AAE, and 38 were classified as non-African American (37 white; 1 Asian-Pacific) and speakers of SWE.

As detailed in Oetting and McDonald (2002), the children's dialect status as an AAE or SWE speaker was confirmed through blind listener judgments and the children's use of 35 different nonmainstream English grammar structures during an elicited language sample. Within each dialect group, half of the children were four years of age (4N) and the others were six (6N); 56% were male and all were considered to be developing language typically.

Table 1 provides available test data from the children, and these included scores from the *Columbia Mental Maturity Scale* (CMMS, Burgmeister, Blum, & Lorge, 1972), the *Peabody Picture Vocabulary Test: Revised* (PPVT-R, Dunn & Dunn, 1981), three subtests that make up the Syntax Quotient of the *Test of Language Development: Primary* (TOLD, Newcomer & Hammill, 1988), and the *Goldman Fristoe Test of Articulation* (GFTA, Goldman & Fristoe, 1986). Twelve children in the 4N group were too young for the normative data of the TOLD; however, data were complete for all of the other tests and on these tests, all of the children earned a standard score above −1 SD of the normative mean. Also, of those who completed the TOLD, all but two earned a standard score above −1 SD (the exceptions were a SWE 4N child who earned a standard score of 79 and a SWE 6N child who earned a score of 83).

Identification and Coding of BE

As part of the original studies, a language sample was elicited from each child by an examiner in a quiet room at the children's schools. Toys and pictures were used as prompts. The samples were then transcribed and coded using *Systematic Analysis of Language Transcripts (SALT)* software (Miller & Chapman, 1992). *SALT* software was also used to extract and further code the children's productions of BE from their language samples. As shown in Table 1, the samples contained 13,443 complete and intelligible utterances and 3,560 code-able BE contexts. Coding of these BE contexts focused on whether the context was overtly marked (*he is walking*) or zero marked (*he ∅ walking*) and as a function of its person, number and tense (*am* vs. *is* vs. *are* vs. *was* vs. *were*), contractibility (+/− whether the form could be contracted in English), and grammatical function (copula vs. auxiliary). To be consistent with other studies, copular and auxiliary BE productions involving *ain't* (e.g., *he ain't gonna do it*) as well as productions of the go copula (e.g., *here go a doll*) and the non-equivalent habitual *be* morpheme (e.g., *he be bad*) were excluded from the analyses.

When coding for the person, number, and tense of the BE form, 94 BE forms involved cases of *was* and *is* leveling. Leveling occurs when a singular form of BE is produced with a plural subject (e.g., Me and my brother was playing…; And two mans was going…; We's gonna …; They is, see?). The AAE group produced 42 (6N = 21; 4N = 21) *was* forms and 7 (6N = 6; 4N = 1) *is* forms, and the SWE group produced 21 (6N = 7; 4N = 14) and 24 (6N = 16; 4N = 8), respectively; 54 served a copular function and 40 served an auxiliary function. When these BE forms were identified in the samples, they were coded based on their surface properties. Thus, the *was* forms in the first two examples above were coded as overtly marked uncontractible auxiliaries, the *is* in the third example was coded as an overtly marked contractible auxiliary, and the *is* in the fourth example was coded as an overtly marked uncontractible copula based on the utterance that preceded it.

When coding for contractibility of BE, it was also necessary to determine how best to code zero marked forms of BE when they occurred within questions. Examples of these included: Y*ou gonna come tomorrow?, That with the boy?, This where you go?*, and *When you coming back*?. When questions included a Wh- form (e.g., *When you coming back*?), the zero marked BE forms were coded as uncontractible (e.g., *When ∅ you coming back*?); however, when a Wh- form was not included, the contractibility of the zero marked forms was ambiguous in some cases because questions in AAE and SWE can be formed with or without inversion (e.g., *Is he coming tomorrow?* which reflects an uncontractible context vs.

He's coming tomorrow? which reflects a contractible context). To determine how best to code these types of BE contexts, we examined all of the children's questions that contained an overtly marked BE form. Although the children produced 1,320 questions within their samples, only 56 (AAE = 19; SWE = 37) contained an overtly marked BE form whose coding would have been ambiguous had it been zero marked. For the AAE groups, 79% of these BE forms were in questions that were not inverted, and for the SWE groups, 97% of these BE forms were in questions that were inverted. Given these findings, we coded the AAE-speaking children's zero marked BE forms within these types of questions as if the questions were not inverted and the SWE-speaking children's zero marked BE forms as if the questions were inverted. Other studies of children and adults have also documented the noninversion of questions in AAE (Green, 2011; Washington & Craig, 2002); however, to our knowledge, similar studies have not been completed with children and adults who speak SWE.

Reliability

Reliability of the language sample transcriptions was examined as part of the original studies and was found to be above 90%. Reliability of the BE coding was examined by having a second research assistant independently code 20% (n = 12) of the samples. Within these samples, there were 680 BE contexts. Inter-rater agreement in the identification and coding of these BE contexts averaged 98% (range across children = 94 – 100%).

Results

Analysis of Variance

In Table 2, the children's percentages of overt marking for BE by each dialect and age group are presented. To examine these data, we first completed a mixed ANOVA with dialect and age as between-subject variables and the three constraints (person/number/tense, contractibility, and grammatical function) as within-subject variables. For the between-subject variables, main effects were significant for dialect, $F(1,58) = 88.42$, $p < .0001$, partial $\eta^2 = .17$. For the within-subject variables, main effects were observed for person, number, and tense, $F(4,191) = 42.06$, $p < .0001$, partial $\eta^2 = .22$, and grammatical function, $F(1,61) = 24.96$, $p < .0001$, partial $\eta^2 = .15$.

Unfortunately, the null effect for contractibility as well as the non-significant interactions between the children's dialects and the three linguistic constraints within the ANOVA are difficult to interpret because the various BE tokens occurred in the different within-subject contexts at labile frequencies that were not controlled. To illustrate, Table 3 presents the total number of BE tokens for all of the children in the AAE 4N group. As can be seen, the lowest number of BE tokens for any given context was 0, and 0 tokens were found for 6 out of the 96 possible cells. Further, the number of tokens per child fluctuated from a low of 23 to a high of 92. Such a highly unbalanced data set is the product of the observational nature of language sample data. Whereas in the aggregate, an ANOVA approach can reasonably assess between-subject differences, it does not accurately assess the effects of within-subject variables such as the linguistic constraints studied here when there is such severe unbalance in the data. The assumption of normality in an ANOVA approach also does not hold with binomial dependent variables (i.e., overt BE vs. zero BE) when the overall proportions for cells are less than 30% or greater than 70%. The relationship, especially outside of this 30 – 70 percent boundary, between the dependent variable and the independent variables is no longer linear, and linearity is a crucial assumption of an ANOVA (see Cedergren & Sankoff, 1974, Sankoff & Labov, 1979, and Sankoff, 1988 for early discussions of these issues as well as more recent work by Quené & van der Berg, 2008, and Tagliamonte, 2006).

Logistic Regression

An alternative approach to data analysis that allows for the binary nature of a dependent variable (i.e., overt BE vs. zero BE) as well as the unbalanced nature of data is logistic regression (see Sankoff, 1988; Tagliamonte, 2006; Walker, 2010 for in-depth discussion). A Generalized Linear Model (GLM) approach to logistic regression is implemented by most modern statistical programs (Hosmer & Lemeshow, 2000). This approach treats the relationship between a binary dependent variable and the independent variables as logistic and transforms the dependent variable accordingly, where a p context represents the probability of a linguistic variant occurring in a particular context:

$$\log\left(\frac{p}{1-p}\right)=m+p_{Context1}+p_{Context2}+\ldots$$

There are two approaches to assessing the statistical significance of a factor (i.e., variable) within a logistic regression analysis. In the first approach, used by many sociolinguists, a step-up and step-down procedure selects a sub-set of statistically significant factors from all factors with a given alpha, such as .05. In the second approach, statistical significance is assessed on each factor individually. Our work here follows the second approach. With this approach, we also employ the comparative method outlined in Poplack and Tagliamonte (2001) and presented in detail by Tagliamonte (2002). As such, we focus on the estimated probabilities for contexts that obtain statistical significance. From these estimated probabilities, we then estimate the strength of each factor's effect on the children's overt marking of BE. In the results discussed below, we use the GENMOD procedure of SAS to produce the results and transform the log-odds into more easily interpretable probabilities.

In Table 4, the results of the multi-factor logistic regression are presented for both dialect groups with the data collapsed for age given that age was not identified as a significant variable in the earlier ANOVA. The factor weights represent the likelihood of the children producing an overtly marked BE form instead of a zero marked form in each context and with respect to the corrected mean (i.e., the estimated average of overt marking across all contexts). Factor weights above .50 are said to favor overt marking of BE whereas weights below .50 are said to disfavour overt marking. However, with the comparative method, the importance of this multi-factor analysis is not the particular numerical value of the weights, but the *ordering* of the weights. The unbalanced nature of conversational data leads to weights that will fluctuate between studies, but the ordering of the weights is more stable.

As can be seen in the table, the results were consistent with the ANOVA approach in that they showed lower rates of overt marking for the AAE group than then SWE group (corrected means for AAE = 64% vs. SWE = 93%). More interestingly, however, were the factor weights associated with the three linguistic constraints. The factor weights indicated that all three linguistic constraints obtained statistical significance for the AAE group: person, number and tense, $\chi^2(3) = 259.96$, $p < .001$, contractibility, $\chi^2(1) = 12.26$, $p > .001$, and grammatical function, $\chi^2(1) = 54.16$, $p < .001$. When considering the effect of person, number and tense of BE, the AAE group were more likely to overtly mark BE in *was/were* contexts (factor weight = .90) than in *am* (factor weight = .84), *is* (factor weight = .22) or *are* (factor weight = .09) contexts.

As can be seen by these results, the factor weights of the logistic regression offer more information about the children's grammars than the raw percentages or the analyses of variance. As an example, the factor weight of .09 for *are* is an estimate of the probability of the children overtly marking *are* while controlling for the effects of contractibility and grammatical function and while also allowing for uneven numbers of tokens across factors.

The raw percentages and the group averages examined within the ANOVA do not allow for generalizations about the magnitude of the *are* effect on the children's marking of BE even though a statistically significant main effect was identified for person, number, and tense within the ANOVA analysis.

The *range* of the weights represents another important component of a multi-factor logistic regression analysis. The range is the distance between the largest weight and the smallest weight, and the range represents the strength (or magnitude) of the effect on the children's overt marking of BE. For person, number and tense, the range for the AAE group was 81 (.90–.09 = 81, with the decimal removed to be consistent with methods used in the field of sociolinguistics). The taxonomy used to interpret the *range* is summarized by Horvath and Horvath (2003): <10 is a weak effect; 10–30 is a moderate effect; 30–50 is a strong effect; >50 is a very strong effect. Using these guidelines, the data showed that the person, number, and tense of the BE form had a very strong effect on the AAE group's overt marking. Moreover, each of the BE contexts obtained statistical significance when tested by post-hoc Wald-Z scores: *was/were*, Z = 3.49, $p = .009$; *am*, Z = 4.17, $p < .0001$; *is*, Z = −7.07, $p < .0001$; and *are*, Z = −8.96, $p < .0001$.

The second variable that obtained statistical significance for the AAE group was the grammatical function of BE. For this variable, copular contexts favored overt marking of BE (factor weight = .64) and auxiliary contexts disfavoured overt marking (factor weight = .36). As indicated by a range of 28, grammatical function showed an overall moderate effect on the AAE group's overt marking of BE.

The third variable to have obtained statistical significance for the AAE group was the contractibility of BE. For this variable, uncontractable contexts favoured overt marking (factor weight = .59) and contractable contexts disfavoured overt marking (factor weight = .41). As indicated by a range of 18, contractibility had an overall moderate effect on the AAE's group's overt marking of BE. Based on these results, the ordering of the factors can be described in a *constraint hierarchy* by their magnitude: person, number and tense > grammatical function > contractibility.

For the SWE group, the person, number and tense of BE obtained statistical significance, $\chi^2(3) = 90.32$, $p < .001$, and the effect was very strong (range = 68). Although the direction of the effect is similar to what was observed for the AAE group: *was/were* > *am* > *is* > *are*, only *are* obtains statistical significance when tested by post-hoc Wald-Z scores: *are*, Z = −8.19, $p < .001$. The SWE child utterances with *are* zero marked were varied as evidenced by the following examples: *Even when you Ø big; We Ø not gonna go; Some of them Ø from my birthday…; Two people Ø on the ground; …didn't bring the toys that Ø the same; They Ø going fishing; What Ø these for*? Thus, the effect of the person, number, and tense constraint on SWE-speaking children's marking of BE, while limited to the form *are*, was robust across various types of utterances and grammatical constructions.

The grammatical function of BE for child SWE also obtained statistical significance, $\chi^2(1) = 41.38$, $p < .001$, and showed a moderate effect (range = 26) on the children's overt marking of BE, with the direction of the effect showing copula > auxiliary. This finding was consistent with the results of the AAE group. Unlike the AAE group, however, the SWE group's rate of overt marking did not show a significant effect for the contractibility of the BE form. For both of these contexts, rates of overt marking were high (93% and 94%).

Discussion

In this study, AAE- and SWE-speaking children's overt marking of BE was examined as a function of their dialect and age, and as a function of three linguistic constraints that have

been shown to influence nonmainstream English-speaking adults' overt marking of BE. From the first set of analyses which made use of ANOVA, the results showed that the children's rates of overtly marked BE varied by their dialect but not their age, with lower rates of overt marking observed for the AAE child speakers than for the SWE child speakers. The person, number, tense, and the grammatical function of the BE forms were also found to influence the children's rates of overt marking within the analyses, but after considering the unbalanced and binary nature of the children's BE productions, the ANOVA approach was deemed limited in its ability to evaluate the influence of the three linguistic constraints.

From the second set of analyses which made use of logistic regression, the results again showed AAE-speaking children to produce a lower rate of overtly marked BE forms than the SWE-speaking children but these analyses also revealed differences in the constraint hierarchies manifested within the two dialects. Whereas the constraint hierarchy for the AAE-speaking children involved all three linguistic variables (person, number and tense > grammatical function > contractibility), the SWE-speaking children's constraint hierarchy was limited to two (person, number and tense > grammatical function). Moreover, although BE marking in both child AAE and child SWE was influenced by the person, number, and tense of BE, the pattern of statistical significance within the dialects varied. For child AAE, the proportions of overtly marked BE forms for *am, is, are*, and *was/were* were all statistically different from the proportion of all forms marked. For child SWE, only *are* was overtly marked at a proportion that statistically differed from (and was lower than) the proportion of all forms marked.

The results documented here for child AAE and child SWE are consistent with the adult nonmainstream English literature. Recall that in adult AAE, all three constraints have been shown to influence rates of overtly marked BE, and this is exactly what was found for the child AAE speakers. The direction and magnitude of the constraints within child AAE also aligns with what has been reported for adult AAE. Like adult AAE, the child AAE speakers produced low rates of overt marking for *are*, higher rates of overt marking for *is*, and high rates of overt marking for *am* and *was/were*. Like adult AAE, the child AAE speakers produced higher rates of overt marking for copular contexts than auxiliary contexts and higher rates of overt marking for uncontractible contexts than for contractible contexts.

In adult SWE, lower rates of overt marking have been consistently documented for *are*, and effects have also been documented for the linguistic constraint of grammatical function, with higher rates of marking for copular contexts than for auxiliary contexts. Like the adult SWE literature, the child SWE speakers produced low rates of overt marking for *are* only and produced higher rates of marking for copular contexts than for auxiliary contexts. Moreover, although the effect for grammatical function was statistically significant, the SWE children's rates of overt marking for both the copular and auxiliary contexts were relatively high (96% and 87%), and this finding is similar to Hazen's (2001) results for SWE-speaking adults (copular =96% vs. auxiliary =78%). Finally and as far as we know, no adult SWE study has revealed an effect for contractibility, and an effect for this constraint was not documented for the SWE-speaking children studied here.

The results presented here were made possible by applying two types of statistical analyses to the data. The traditional ANOVA method allowed us to detect important quantitative differences in the children's rates of overt marking, and logistic regression allowed us to uncover important qualitative differences in the children's constraint hierarchies. When compared to adult AAE and SWE, the results show a high degree of consistency across generations of AAE and SWE speakers. This finding occurred even though none of the adult studies reviewed was conducted in or near the dialect communities of the children.

Moreover, the earliest adult studies reviewed were published in 1969, which was ~30 years before the child data were collected. These results indicate that BE marking, like verbal –s marking but not relative clause marking, is a grammatical structure that is remaining stable across different generations and different communities of AAE and SWE speakers.

Our results for child AAE and child SWE differ from each other and they also contrast with literature on children who speak MAE. When the three dialects are compared, the most striking difference between them involves the person, number, and tense of the BE form. For this linguistic constraint, AAE shows effects for all forms, SWE shows effects for *are* only, and MAE shows no effects. Contractibility also shows interesting dialect differences between AAE, SWE, and MAE. For this constraint, AAE shows effects with the highest rates of overt marking for uncontractible contexts, SWE shows no effects, and MAE, when effects are documented, show the highest rates of overt marking for contractible contexts. Thus, for this linguistic constraint, the presence and direction of the effect differs across the three dialects. For grammatical function, all three dialects are similar to each other and show higher rates of overt marking for copular contexts than for auxiliary contexts.

Although the dialect differences identified in the data are consistent with previous studies, we were unable to detect a significant effect for age in the children's marking of BE. In fact, within each dialect, the average percentages of overt marking were near identical for the four- and six-year-olds (AAE: 58% and 64%; SWE: 87% and 89%). These percentages also appear to align with rates reported in the adult AAE and SWE literature (although overall rates of overt BE marking are not routinely reported for adults). If this finding replicates, then it may be the case that children who speak AAE and SWE reach adult levels of use for BE at an earlier age than children who speak MAE. Additional studies are needed to test this hypothesis. Our study was limited to data from 62 children whose language samples were collected as part of a set of previous studies. The standard deviations that accompanied the group averages of the four- and six-year-olds were also large, especially for the AAE group at 44% and 39%, respectively. In our on-going work, we are systematically recruiting larger numbers of AAE- and SWE-speaking children who vary in their maternal education levels and test scores, and in the rates at which they produce nonmainstream grammar structures to determine if these child variables can explain some of the within-dialect variability that occurs in studies of nonmainstream English speakers. For BE, the findings from the current work indicate that any future study of these types of child attribute variables must control for the form's person, number, and tense, contractibility, and grammatical function.

Speech-language clinicians should be able to use the current set of findings to better understand and articulate the ways in which children's grammars vary across different dialects of English. As shown here, the dialect variation involves differences in the rates of overt marking and in the linguistic constraint hierarchies. In other words, these child dialects differ in the rates at which they overtly mark BE and in their linguistic conditioning of BE, the latter of which influences when a particular type of marking (zero or overt) will be produced.

The findings also have implications for assessment, treatment of children with language impairments, and consultative services to teachers of nonmainstream English-speaking children without language impairments. For assessment, the findings can be used to advocate for language evaluations to include a wide range of grammar structures. For BE, this evaluation should target multiple forms and contexts to fully capture children's rates of use and their dialect-specific constraint hierarchies. For treatment of children with language impairments and whose rates of BE marking and constraint hierarchies are shown to fall outside of a dialect-appropriate range, the findings can be used to strategically target some forms and contexts of BE before others.

Finally, for consultative services, the findings can be used by speech-language clinicians to evaluate the test and teaching materials of a classroom. For BE, the current findings suggest that classroom materials should not present a limited number and type of BE forms. BE form variability in classroom materials should help nonmainstream English-speaking children demonstrate their knowledge when they are tested. BE form variability in classroom materials should also help these same children expand their understanding and use of grammar in school and when interacting with same dialect- and different dialect-speaking communication partners.

Acknowledgments

This study was made possible by RO3 DC03609 and RO1 DC009811 from the National Institute on Deafness and Other Communication Disorders. The authors would like to thank the children, parents, and teachers who made data collection possible, Tina Villa for completing an analysis of the children's questions, and Nicté Fuller Medina, Kyomi Gregory, Allison Lealess, Ryan Lee, Jessica Richardson, Andrew Riviére, Christy Seidel, and Tina Villa for providing valuable feedback on an earlier draft of this manuscript.

References

Bailey G, Maynor N. The present tense of be in southern Black folk speech. American Speech. 1985; 60:195–213.

Baugh, J. A re-examination of the Black English copula. In: Labov, W., editor. Locating language in time and space. Academic; New York: 1980. p. 83-195.

Blake R. Defining the envelope of linguistic variation: The case of “don't count” forms in the copula analysis of African American Vernacular English. Language Variation and Change. 1997; 9:57–79.

Brown, R. A First Language: The Early Stages. Harvard University Press; Cambridge, MA: 1973.

Burgmeister, B.; Blum, H.; Lorge, I. Columbia Mental Maturity Scale. The Psychological Corporation; San Antonio, TX: 1972.

Cedegren H, Sankoff D. Variable Rules: Performance as a Statistical Reflection of Competence. Language. 1974; 50:333–355.

Childs B, Malinson C. African American English in Appalachia: Dialect accommodation and substrate influence. English World-Wide. 2004; 25:27–50.

Cleave PL, Rice ML. An examination of the morpheme BE in children with specific language impairment: The role of contractibility and grammatical form class. Journal of Speech, Language, and Hearing Research. 1997; 40:480–492.

Cleveland, LH. Unpublished doctoral dissertation. Louisiana State University; 2009. Children's productions of verbal –s by dialect type and clinical status.

Cukor-Avila, P. Co-existing grammars: The relationship between the evolution of African American and White Vernacular English in the South (-). In: Lanehart, S., editor. Social and Historical Contexts of African American English. John Benjamins; Amersterdam: 2001.

de Villiers JG, de Villiers PA. A cross-linguistic study of the acquisition of grammatical morphemes. Journal of Psycholinguistic Research. 1973; 2:267–278.

Dubois S, Horvath BM. From accent to marker in Cajun English: A study of dialect formation in progress. English World-Wide. 1998; 19:161–188.

Dubois S, Horvath BM. Let's tink about dat: Interdental fricatives in Cajun English. Language Variation and Change. 1999; 10:245–261.

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test - Revised. American Guidance Service; Circle Pines, MN: 1981.

Fasold, R.; Wolfram, W. Some linguistic features of Negro dialect. In: Fasold, R.; Shuy, R., editors. Teaching Standard English in the Inner City. Center for Applied Linguistics; Washington, DC: 1970. p. 41-86.

Goldman, R.; Fristoe, M. Goldman-Fristoe Test of Articulation. American Guidance Service; Circle Pines, MN: 1986.

Garrity A, Oetting J. Auxiliary BE production by African American English-speaking children with and without specific language impairment. Journal of Speech, Language, and Hearing Research. 2010; 53:1307–1320.

Green, LJ. Language and the African American child. Cambridge University Press; Cambridge, MA: 2011.

Hazen, K. Identity and Ethnicity in the Rural South: A sociolinguistic view through past and present BE. Duke University Press; Durham, North Carolina: 2001.

Harris, R.; Zimmerman, J. The Southern Rural Development Center Policy. Mississippi State, MS: 2003. Children and poverty in the rural South. Series, No. 2

Hollingshead, A. Four-factor index of social position. Yale University; 1975. Unpublished manuscript

Horton-Ikard R, Ellis Weismer S. Distinguishing African-American English from developmental errors in the language production of toddlers. Applied Psycholinguistics. 2005; 26:597–620.

Horvath BM, Horvath RJ. A closer look at the constraint hierarchy: Order, contrast and geographic scale. Language, Variation and Change. 2003; 15:143–170.

Hosmer, DW.; Lemeshow, S. Applied Logistic Regression. 2nd Edition. John Wiley & Sons, Inc; New York: 2000.

Jackson S, Roberts J. Complex syntax production of African American preschoolers. Journal of Speech, Language, and Hearing Research. 2001; 44:1083–1096.

James SL, Kahn LM. Grammatical morpheme acquisition: An approximately invariant order? Journal of Psycholinguistic Research. 1982; 11:381–388.

Johnston, JR.; Schery, TK. The use of grammatical morphemes by children with communication disorders. In: Morehead, D.; Morehead, A., editors. Normal and deviant child language. University Park Press; Baltimore, MD: 1976. p. 239-258.

Labov W. Contraction, deletion, and the inherent variability of the English copula. Language. 1969; 45:715–762.

Labov, W. Language in the inner city. University of Pennsylvania Press; Philadelphia, PA: 1972.

Leonard LB, Deevy P, Miller CA, Charest M, Kurtz R, Rauf L. The use of grammatical morphemes reflecting aspect and modality by children with specific language impairment. Journal of Child Language. 2003; 30:769–796. [PubMed: 14686084]

Miller, J.; Chapman, R. Systematic Analysis of Language Transcripts [Computer software]. Language Analysis Laboratory; University of Wisconsin; Madison: 1992.

Miller, K.; Weber, B. Persistent poverty and place: How do persistent poverty dynamics and demographics vary across the rural urban continuum. 2004. Retrieved from the Southern Rural Development Center: http://srdc.msstate.edu

Newcomer, P.; Hammill, D. Test of Language Development–Primary. Second Edition. ProEd; Austin, TX: 1988.

Oetting JB, Garrity AW. Variation within dialects: A case of Cajun/Creole influence within child AAE and SWE. Journal of Speech, Language and Hearing Research. 2006; 49:16–26.

Oetting JB, McDonald JL. Nonmainstream dialect use and specific language impairment. Journal of Speech, Language, and Hearing Research. 2001; 44:207–223.

Oetting JB, McDonald JL. Methods for characterizing participants' nonmainstream dialect use in child language research. Journal of Speech, Language, and Hearing Research. 2002; 45:505–518.

Oetting JB, Newkirk BL. Subject relatives by children with and without SLI across different dialects of English. Clinical Linguistics and Phonetics. 2008; 22:111–125. [PubMed: 18253870]

Oetting JB, Newkirk BL. Children's relative clause markers in two nonmainstream dialects of English. Clinical Linguistics and Phonetics. 2011; 25:725–740. [PubMed: 21453062]

Oetting JB, Pruitt SL. Use of Southern African American English across groups. International Journal of Multicultural Communication Disorders. 2005; 3:136–144.

Polite E, Leonard LB. A method for assessing the use of first person verb forms by preschool-aged children with SLI. Child Language Teaching and Therapy. 2007; 23:353–366.

Poplack, S.; Tagliamonte, S. African American English in the Diaspora. Blackwell Publishers; Oxford: 2001.

Quené H, van den Bergh H. Examples of mixed-effects modeling with crossed random effects and with binomial data. Journal of Memory and Language. 2008; 59:413–425.

Rice ML, Wexler K, Hershberger S. Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. Journal of Speech, Language and Hearing Research. 1998; 41:1412–1431.

Rickford, J. The creole origin of African American Vernacular English: Evidence from copula absence. In: Mufwene, S.; Rickford, J.; Bailey, G.; Baugh, J., editors. African American Vernacular English Structure, History and Use. Routeledge; New York: 1998. p. 154-200.

Rickford, JR. African American Vernacular English. Blackwell; Malden, MA: 1999.

Rickford J, Ball A, Blake R, Jackson R, Martin N. Rappin on the copula coffin: Theoretical and methodological issues in the analysis of copula variation in African-American Vernacular English. Language Variation and Change. 1991; 3:103–132.

Romaine, S. Socio-historical linguistics: Its status and methodology. Cambridge University Press; Cambridge: 1982.

Sankoff, D. Variable Rules. In: von Ulrich Ammon, Herausgegeben; Dittmar, Norbert; Mattheier, Klaus J., editors. An International Handbook of the Science of Language and Society. Mouton De Gruyter; Berlin: 1988. p. 984-996.

Sankoff D, Labov W. On the uses of variable rules. Language in Society. 1979; 8:189–222.

Sankoff D, Tagliamonte S, Smith E. GOLDVARB X. 2005

Tagliamonte, S. Comparative sociolinguistics. In: Chambers, JK.; Trudgill, P.; Schilling-Estes, N., editors. The Handbook of Language Variation and Change. Blackwell; Oxford: 2002. p. 729-763.

Tagliamonte, S. Analyzing Sociolinguistic Variation. Cambridge University Press; Cambridge: 2006.

Walker, J. Rephrasing the copula: Correction and zero in early African American English. In: Poplack, S., editor. The English History of African American English. Blackwell; Oxford: 2000. p. 35-72.

Walker, J. Variation in Linguistics Systems. Routledge; New York: 2010.

Washington JA, Craig HK. Dialectal forms during discourse of poor, urban, African American preschoolers. Journal of Speech and Hearing Research. 1994; 37:816–823. [PubMed: 7967568]

Washington JA, Craig HK. Morphosyntactic forms of African American English use by young children and their craegivers. Applied Psycholinguistics. 2002; 23:209–231.

Winford D. Another look at the copula in Black English and Caribbean Creoles. American Speech. 1992; 67:21–60.

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

Wolfram W. The relationship of White Southern speech to vernacular Black English. Language. 1974; 50:850–872.

Wolfram, W.; Schilling-Estes, N. American English: Dialects and variation. Blackwell; Malden, MA: 1998.

Wolfram, W.; Thomas, E. The development of African American English. Blackwell Publishing; Malden, MA: 2002.

Wyatt, T. Unpublished dissertation. University of Massachusetts; 1991. Linguistic constraints on copula production in Black English child speech.

Wyatt, T. The acquisition of African American English copula. In: Kamhi, A.; Pollock, K.; Harris, J., editors. Communication development and disorders in African American children. Paul H. Brookes Publishing Co; Baltimore, MR: 1996. p. 95-116.

Table 1

Participant characteristics and description of language samples

	AAE		SWE	
	4N	**6N**	**4N**	**6N**
N	12	12	19	19
Males	5	5	12	13
Ages in Months	56.83 (3)	74.5 (4)	48.3 (5)	76.1 (6)
CMMS[a]	102.4 (7)	101.3 (4)	104.7 (10)	104.3 (9)
PPVT[b]	97.8 (8)	102.2 (13)	102.2 (7)	104.9 (12)
TOLD[c]	97.08 (11)	101.83 (11)	96.29 (7.4)	104.8 (13)
GFTA[d]	80.1 (15)	95.6 (6)	88.8 (8)	92.8 (13)
Mean C & I Utterances[e]	192 (63)	221 (79)	204 (38)	242 (47)
Total C& I Utterances[f]	2302	2652	3876	4613
Total BE Tokens[g]	559	693	1059	1249

[a] *Columbia Mental Maturity Scale*, M = 100, SD = 15;

[b] *Peabody Picture Vocabulary Test: Revised*, M = 100, SD = 15;

[c] Syntactic quotient from the *Test of Language Development Primary*, M = 100, SD = 15;

[d] *Goldman Fristoe Test of Articulation*, percentile scores reported;

[e] Mean number of complete and intelligible utterances in each child's sample;

[f] Total number of complete and intelligible utterances in samples;

[g] Total number of BE tokens coded for the constraints of interest.

Table 2

Percent of overt marking of BE.

	AAE		SWE	
	4N	**6N**	**4N**	**6N**
Person/Number				
Am	100 (0)	91 (26)	95 (14)	96 (18)
Is	38 (31)	57 (32)	88 (24)	94 (19)
Are	34 (38)	31 (40)	72 (41)	75 (38)
Was/were	94 (17)	94 (19)	97 (17)	99 (6)
Contractibility				
Contractible	56 (39)	53 (40)	86 (27)	88 (26)
Uncontractible	62 (42)	76 (36)	88 (31)	91 (26)
Grammatical Function				
Copular	62 (37)	72 (34)	95 (14)	91 (24)
Auxiliary	55 (43)	57 (43)	76 (38)	88 (29)
Total	58 (44)	64 (39)	87 (29)	89 (26)

Table 3

Frequency of BE contexts for 4N AAE group.

Participant Identification Number												
	100	101	103	105	110	112	114	115	117	118	119	120
Total BE Tokens												
	47	26	38	23	38	62	38	92	49	26	94	25
BE Tokens by Person, Number, Tense												
Am	7	0	0	5	2	2	0	4	3	4	2	0
Is	34	20	21	9	15	39	23	67	25	19	57	12
Are	2	0	5	3	3	13	6	13	12	3	24	2
Was/were	4	6	12	6	18	8	9	8	9	0	11	11
BE Tokens by Contractibility												
Contractible	41	16	26	13	20	37	25	64	30	18	69	11
Uncontractible	6	10	12	10	18	25	13	28	19	8	25	14
BE Tokens by Grammatical Function												
Copular	28	19	24	18	22	46	24	64	36	20	44	11
Auxiliary	19	7	14	5	16	16	14	28	13	6	50	14

Table 4

Logistic regression for the selection of overt BE for AAE and SWE[a].

	AAE			SWE		
	FW	**%**	**N**	**FW**	**%**	**N**
Overall Rate	.78	64	1248	.95	93	2294
Person, Number, Tense						
Are	.09	27	175	.16	77	192
Is	.22	59	740	.41	95	1518
Am	.84	94	67	.59	96	192
Was/were	.90	96	263	.84	99	319
Range	81			68		
Contractibility						
Contractable	.41	57	818	[]	93	1562
Uncontractable	.59	77	430	[]	94	578
Range	18			--		
Grammatical Function						
Copular	.64	70	807	.63	96	1571
Auxiliary	.36	53	441	.37	87	569
Range	28			26		

[a]FW = factor weight. The weight reported reflects the estimated probability for the overall rate of overt BE from the logistic regression. The % reflects the proportion of BE contexts overtly marked for the group as a whole. The N reflects the number of BE contexts included in the analysis.