

Copula Variability in Jamaican Creole and African American Vernacular English: A Reanalysis of DeCamp's Texts*

John R. Rickford
Stanford University

1. Introduction

The purpose of this paper is to reanalyze copula variability in the four Anansi story texts in DeCamp (1960), the classic Jamaican Creole (JC) data-set which Holm (1976, 1984) first used to show synchronic parallelisms, and therefore potential diachronic links, between creoles and African American Vernacular English (AAVE). To anticipate my major finding: After reanalysis, the quantitative patterns of copula absence by following syntactic environment in JC turn out to be much more similar to those in AAVE, lending even further weight to the hypothesis that AAVE is a decreolized form of an earlier plantation creole that was typologically similar to JC.

Before turning to the substantive issues, I wish to make a few remarks about why I chose to submit this particular paper for a volume honoring William Labov. There is, first of all, the fact that Labov's (1969) analysis of the AAVE copula remains a high point of his career; key elements of that analysis (for instance, the regular relation between contraction and deletion) are familiar to very many linguists, within sociolinguistics as well as other subfields, and the analysis itself introduced the variable rule framework, which remains quite central within variation theory. Secondly, the creole origins hypothesis — particularly as affected by the similarity between copula absence in AAVE and various Caribbean creoles — is one Labov addresses in several major publications, including Labov (1972: 36-64) and Labov (1982). Thirdly, this paper involves quantitative analysis of data drawn from recorded samples of natural speech, adhering scrupulously to Labov's important (1969:737, fn. 20) principle of accountability. Finally, while focussing on

specific analytical problems, it illustrates Labov's general point (1969:728) that decisions about what to count and how to define the envelope of variation for a linguistic variable pose "subtle and difficult problems;" for the variationist, but also crucial and substantive ones.

Copula variability in Jamaican Creole (JC) has played a major role in debates about the history of African American Vernacular English (AAVE). In particular, the fact that JC (and Gullah) displayed a pattern familiar from studies of AAVE — more copula absence before adjectives than before localatives and noun phrases — was taken as strong evidence for the creole ancestry of the latter by Holm (1976, 1984), Baugh (1979, 1980) and Labov (1982). Figure 1, drawing on data from Jamaica, NYC and LA, shows the high adj/low NP and locative pattern which John Sleded regarded as "the first serious evidence for the creole hypothesis" (Labov 1982:198, fn. 26).

However, as further study of Figure 1 will confirm, the Jamaican data were also an embarrassment for the creole hypothesis in showing relatively low rates of zero copula before *Verb(+ing)* and *goma*, where AAVE typically showed its HIGHEST rates of *is/are* absence. Except for Holm (1984:293-4), no one really drew attention to this disparity, but it bothered me for years, since low zero copula rates before *Verb(+ing)* and *goma* did not

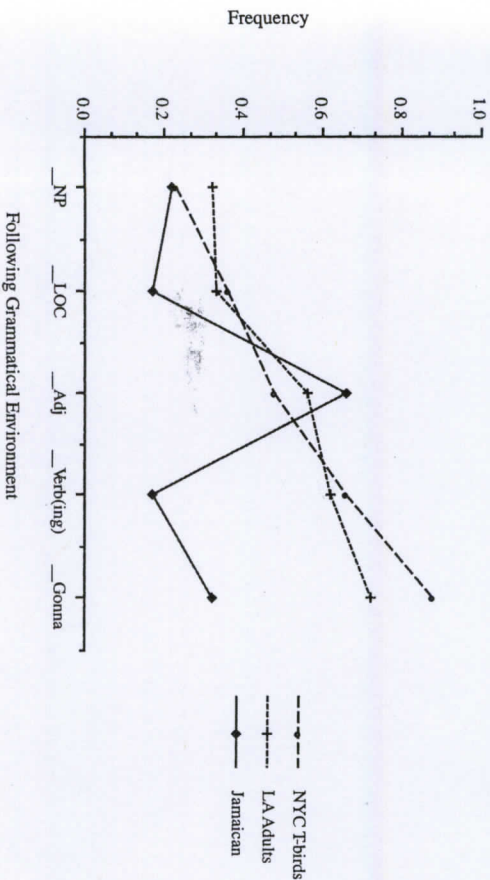


Figure 1. Copula Absence in 3 African-American Dialects

accord with my intuitions as a native speaker of Guyanese Creole (GC), nor with my research experience with Gullah, spoken on the South Carolina Sea Islands.

In a 1990 paper on Barbadian (first presented at the 1988 NWAV meeting in Montreal), Renee Blake and I suggested that the low zero figures for JC *Verb(+ing)* and *goma* might have resulted from the way in which the envelope of variants was drawn in Holm's original (1976, 1984) analysis of JC.¹ As we noted (Rickford & Blake 1990:261):

In the *Verb(+ing)* case, for instance, only \emptyset and inflected *is* or *are* can occur in equivalent syntactic slots; basilectal *de* and *a* cannot co-occur with *Verb(+ing)* (*"dem de wakin") but only with *Verb* ("dem de go") and therefore tokens with these variants should not be considered along with the others. However, 82% of the variants in Holm's preverbal subcategory for Jamaican come from *de* and *a*: if these are removed, leaving only tokens of inflected *be* and \emptyset , the proportion of zero for *Verb(+ing)* climbs to 89%... A similar categorization or computation error probably accounts for the low *goma* figure which Holm reports for Jamaica. (Rickford and Blake 1990:261)

This paper is essentially an amplification of this suggestion, but one in which I have recorded and tabulated EVERY copula variant in the DeCamp (1960) data set that Holm (1976) first analyzed. One reason for doing this is to recover the full information on sample sizes and 'don't count' tokens per subcategory which Holm's analysis of relative frequencies — though valuable — did not provide. Another reason is to see whether other factors besides following grammatical environment might be involved. I will present VARBRUL analyses of these towards the end of my paper. A final reason is to have this reanalysis of DeCamp's texts serve as a pilot for the analysis of new data which I recorded in March 1991 in Jamaica and Barbados, as part of a larger study of copula variability in AAVE, the creoles, and other languages.

Before presenting any new analysis, however, I wish to say some more about the JC data set being used in this paper. In 1960 David DeCamp published phonetic transcripts and translations of four stories which he had recorded in Accompong village, a former Maroon stronghold, in Jamaica. The stories, ranging in length from about a quarter of an hour to half an hour each, were told by Mr. Emmanuel "Baba" Rowe, who was nearly 80 years old at the time. Although the stories contain some elevated, Standard English (SE) elements which probably represent Mr Rowe's adjustments to his high-status American interlocutor, they are dramatically delivered, and replete with

basilectal or "deep creole" elements (See B Bailey 1971, Holm 1984). The following brief extract from one of his stories — one entitled "Andrew and the Old Witch" — will give an idea of what DeCamp's stories, transcripts and glosses are like:

- (1) From Emmanuel Rowe's JC story "Andrew and the Old Witch" (DeCamp 1960: 159): (Text, accents and gloss as in original; but copula tokens which were counted in my analysis are indicated in boldface.)
- di uol 'liedi 'tiek op in 'rieza. him 'waip di 'rieza, 'wen in 'de go
'shaapin di 'rieza. a 'so in 'shaapin. in 'waip di 'rieza. hin sie,
'shaapin mi 'rieza! 'shaapin mi 'rieza! 'shakam! 'shii! 'shakam! di
'rieza Ø 'shaap 'tel if a 'f'ai 'pich 'pan i, i 'kot im. so wen in 'draa 'op
'tuu di 'bed fi go 'kot di 'gyal 'truot, 'andro 'nuo 'wa 'hin de go. 'andro
'sing 'out 'andro 'tan de. 'luk 'pan im, an 'sii wa de 'don. 'andro sie,
'ool mi 'madam 'kwii 'nan 'ool!
'wiel 'sali 'ool mi 'jienan 'ei!
'wai yoi! 'an mi 'niem Ø 'andro!

Gloss: "The old lady take up her razor. She wipe the razor, when she is going to sharpen the razor. It is so that she sharpens. She wipe the razor. She say, 'Sharpen, my razor! Sharpen, my razor! Shakam! Shee! Shakam!' [merely imitative syllables ... DDeC.] The razor sharp until if a fly pitch upon it, it cut him. So when she draw up to the bed, to go cut the girls' throats, Andrew know what she is going to do. Andrew sing out. Andrew stand there, look upon her and see what is being done. Andrew say,

'Oh, my Madame Queen Anne, oh!
Wake, Sally, oh! My dear Jane Anne, hey!
And my name Andrew.'"

2. Holm's (1976, 1984) Analysis

The beauty of DeCamp's (1960) data set is that his transcripts are publicly available for inspection and reanalysis, and that copies of his original recording are also available — a rarity in our field. The first person to code and analyze copula tokens from this data set in terms of following grammatical environment was John Holm, who published a preliminary analysis in 1976 and a more detailed discussion in 1984. Holm's later paper does not say anything about copula tokens which he had to exclude from consideration, but

Table 1. Copula variants by following grammatical environment in Decamp's JC texts (Holm's analysis)

Variant	_Noun Phr	_Locative	_Adjective	_Verb	_Gonna
Ø	22%	17%	66%	17%	32%
a	31%	—	9%	6%	—
de	0%	45%	2%	76%	— (68%?)
be	47%	17%	23%	2%	—
ben	—	17%	—	—	—

Notes: N=323; adapted from Holm 1984:292 (Table 1) and 293 (Table 2); *be* here and in subsequent tables includes conjugated forms (*iz, ar, waz, wor*).

since his analysis is based on only 343 copula tokens whereas nearly 500 copula slots occur in the transcripts, he clearly did exclude some 'don't count' cases, as everyone does when doing a variation analysis. Holm (1976) gives some indication of what those exclusions were: passives, cases with expletive *there*, and modals ("They are to arrive").

Table 1 shows the relative frequency of copula variants by following environment which Holm (1984) found among the tokens he considered. The percentages in the 'zero' row are of course the data points for the Jamaican line in Figure 1, but the other variants are also critical, and we will see how their percentages match up with our reanalysis below. Holm's analysis of DeCamp's data was extremely valuable in its own right, and insofar as it provided a basis for comparison in Baugh's significant (1979, 1980) reanalysis of Labov et al.'s NYC data and his own LA data, which further strengthened the argument for a creole origin for AAVE.

3. My Reanalysis

Following on some preliminary work on DeCamp's texts by my students², I undertook to reexamine every potential copula token in the transcripts myself — coding each token in terms of following grammatical environment, person and subject type, tense, and other factors. In each case, however, my codings were based on the recording rather than the transcript, which, although generally accurate, does contain occasional errors. I should add that I also numbered every line of the phonetic transcript sequentially from 1 to 842, and that these are the numbers I will use to identify examples employed in this paper, should anyone wish to return to the transcripts to study them in context.

Table 2a. Frequency of "don't count" types (my reanalysis)

DC type	Example	Freq.
Highlighter/left (HD)	a de di haks-dem wok ("It was there ...") (L266)	54
No overt subject (NS)	wen Ø going down tuu di goli, in go up (L43-4)	31
Anterior <i>ben</i> (ben)	di uol lie di no ben hie se ... (L198)	17
Clause final (CF)	wat di ting iz, (L151)	9
Unclear (UN)	wen mi du huom (=de? L477-8)	7
Future w/o <i>go</i> , <i>gwain</i> (FU)	a Ø kil yu tidie (L225)	3
Miscell foll envir (MISC)	dem a fi-hyar fiidn (L423)	4
Incomplete (IN)	yu naa iihm de ... (L676)	2

Table 2b. Copula variant realizations of "don't count" tokens by following grammatical environment (my reanalysis)

Variant	_NP	_Loc	_Adj	_V(+ed)	_V+ing	_V+con	_go V	_Other	Total
Ø	0	0	5	11	1	0	0	1	18
a	34	5	1	0	0	1	0	16	57
de	0	4	0	0	0	13	2	9	28
be(=iz)	3	0	1	0	0	0	0	3	7
ben	0	8	0	7	0	0	0	2	17
TOTAL	37	17	7	18	1	14	2	31	127

One of the first analytical issues I had to deal with is deciding which copula tokens to set aside as 'don't count' (DC) cases, either because they were difficult to classify reliably (indeterminate) or because they behaved more categorically and less variably than the tokens I was going to count and analyze in detail.³ Unlike Holm, I did not exclude the eight existentials in the text, which occur with both full and contracted tokens of conjugated *be*, as in "der iz a drai golf" (36) and "das a neks stuori" (284). Nor did I exclude passives, which occur with full, contracted and zero realizations of English *be*. Nor did I find copulas preceding modals, unless we count as modal the *fi*-phrase in "wa Ø fi don tide" (687), which I counted as a DC-Miscellaneous case. In general my DC tokens were ones in which one or more variants did not or could not occur. Table 2a shows the various DC types, and their frequency in my data, beginning with the most frequent category of exclusions — tokens of highlighter or focussing *a* (in one case *iz*), as in "a de di haks-dem wok" (Line 266).⁴ Table 2b shows the frequency of copula variants among the DC tokens by following grammatical environment. In all, there were 127 DC cases in the texts, leaving a total of 368 'count' cases, just 25

Copula Variability in JC and AAVE

363

Table 3. Copula variants by following grammatical environment in Decamp's texts (my 1st reanalysis)

Variant	_NP n=68	_Loc n=40	_Adj n=48	_V(+ed) n=34	_V+ing n=21	_V+con n=85	_gwain V n=25	_go V n=47
Ø	28%	18%	81%	76%	86%	0%	100%	0%
a	18%	0%	0%	3%	0%	12%	0%	2%
de	0%	65%	0%	0%	5%	88%	0%	98%
be	54%	18%	18%	21%	9%	0%	0%	0%

N=368, excluding 127 "Don't Count" cases

Table 4. Examples of copula variants in each environment

_NP	shii iz a uol wich (132); mii Ø kwaku (712); mi a kwaku (749)
_Loc	dem waz der (353); hin de pan di trii (194); mi Ø down a kam trii (638)
_Adj	shii iz def (147); him Ø def (202); di trii gyol wor jobial (9); in Ø so big (808)
_V(+ed)	di trii Ø kot (221); babiapuo a ded (839); kwaku waz haiti rekomendid (839a)
_V+ing	we Ø gwain at nou (1); him de digin a di kantri (666); dem wo taaking (173)
_V [+contin]	dat tida gyal de kom ya (16); it a bwail (24); andro de shlip (459)
_gwain (tu) V	yu Ø gwain fain out (12); a Ø gwain tu get (438); hin Ø gwain kil dem (326a)
_go V	mi de go tel yu nou (49); a de im a go tan (643a); hin de go signal kwaku (768)

more than Holm found. In the rest of the paper, I will concentrate on these 368 'count' tokens.

Table 3 shows the relative frequency of copula variants in the eight subcategories I think it is necessary to recognize, and Table 4 provides examples of each subcategory. As Table 2b indicates, there is also a small ninth miscellaneous subcategory, for adverbs and so on, but we can safely ignore it from this point on. Beginning with NP, I will now go on to comment on the results in each subcategory and how they compare with Holm's.

3.1 *Noun Phrase*

The results in the __NP column are not too surprising, and not very different from Holm's results (28% zero versus 22%, see Table 1), although it is impossible to do a chi-square comparison to confirm this because we don't have the column Ns for Holm's data. The biggest difference is the higher percentage he reports for the creole nominal copula *a* (31% vs my 18%). However, the __NP column in my Table 2b suggests where the difference may lie: in his recognition of at least some of the thirty-four tokens of

highlighter *a* before NP which I set aside because they have a different semantic function (topicalizing, equivalent to "It is" rather than "is") and are not usually replaceable by \emptyset .

It should also be noted that eight of the 19 NP tokens which make up the 28% figure for zero in Table 3 involve *niem*, as in the last line of the JC text above: "an mi niem \emptyset andro." The justification for analyzing these as \emptyset followed by NP is a single variant with *a* before NP: "mi niem a andro" (line 1896 — on tape but not in the transcript). But these cases could also be analyzed as \emptyset followed by verbal *niem* (which Bailey 1966 treated as a special 'naming' verb, in contrast with the equating verb *a* and the locative verb *de*), or as "Don't Count" (unclear analysis) cases. If, on either count, they were removed from the NP pool (as I now think they should be), the relative frequency of zero copula before ___NP would drop from 28% (19/68) to 18% (11/60), a figure even lower than Holm's.

The larger point — of a piece with the distinctions between 'Labov Deletion,' 'Straight Deletion' and so on which were introduced in Rickford et al. (1991) — is that the classification and counting decisions which lie behind the statistics that variationists present and publish are sometimes problematic, and that how we solve the problems posed by our data crucially affects what we find.

3.2 ___Locative

In the case of the locative column in Table 3, my 18% figure for copula absence (\emptyset) is virtually identical to Holm's corresponding figure of 17%. The major difference between our analyses of the distribution of variants in this subcategory is that he includes *ben*, while I do not. From Table 2b, we see that there are eight "DC" tokens of *ben* with locatives; if we added these to the sample of forty 'count' tokens in the locative subcategory (Table 3) and recalculated the relative frequencies of the prelocative variants accordingly, the relative frequency of *de* would drop to 54% (closer to his 45% in Table 1), and *ben* would account for 17% of the tokens in this subcategory (identical to his 17% in Table 1).

However, one sound reason for excluding *ben* tokens was given by Holm himself (1984:303, fn. 3): that it marks anterior tense and has little to do with the copula beyond its etymology.⁵ This is clearest in the case of the seven verb stems which *ben* precedes in these texts, which translate into Standard English

equivalents with *did* V, *had* V or V+*ed* rather than (copulative) *was* V+*ed*, and have the textual anterior or past-before-the-past semantics (especially with non-statives) associated with creole *bin* (Bickerton 1975:28-29, 46-47; Rickford 1987:137-43) as in this example from DeCamp's fourth text, the story of "Babiabou":

- (2) mi no ben tel yu sie yu mos tan todi an wach wa de go haphm? (line 837)

"Didn't I tell you that you must stand steady and watch what was going to happen?"

For the seven locatives preceded by *ben*, the case is somewhat less clear-cut, since these sometimes do vary with and translate into copulative and simple past *waz*, as in:

- (3) we unu ben de? (line 277) "Where were you guys?" (Compare 'dem waz der,' L 353)

But in other cases, *ben* clearly does not vary with copulative *waz*, carrying instead an anterior preverbal sense similar to that in (2):

- (4) an di wata mount di gyal siem plies we im ben de, ... (lines 118-19)
"And the water mounted the girl to the same place where it had been (previously)"

Six of the seven prelocatives with *ben* in these texts in fact precede a form *de* which seems to function unambiguously as a locative verb (locative copula) rather than adverb (i.e. in example (3) the tensed locative copula is *ben de* and the locative adverb is *we*), making them simply the past tense equivalent of the primary copula variant in this subcategory, non-tensed *de*, as in:

- (5) wen mi de a huom ... (line 371) "When I am at home ..."

If they were to be counted as copula tokens (which is not the analysis I favor at present), they might be better considered as tokens of *de* or at least *bin* *de* rather than *bin*.

One very important point to note about the distribution of the copula variants in the Jamaican data is that although the percent of zero copula for locatives is lower than that for ___NP — something virtually unparalleled in studies of African American Vernacular English — it is the persistence of the

creole copula *de* (65% in Table 3) rather than *be* that is responsible for the low rate of copula absence before locative in the Jamaican data. A comparison of the *be* percentages for NP and Loc in the Jamaican data (54% *be* for NP, 18% *be* for Loc, Table 3) replicates the relationship between the two environments that is generally found in AAVE (more *be* with noun phrases than with locatives), where the *be* percentage is simply the reciprocal (all full and contracted forms) of the percentage of copula absence. Compare, for instance, these relative frequencies which Labov (1969:732, Table 2) reported for the NYC Jets: 68% *be* with NP, 48% *be* with Loc. As Bickerton (1972:651-52) showed in his study of Guyanese Creole, *de* is the most persistent of the creole copulas. It must be the variable redistribution of this persistent, high frequency form to \emptyset and *be* later in the decreolization process which produces the fluctuations in the locative/adjective ordering that many studies of New World African English report (see Rickford et al 1991:121) and which Singler (1991) also reports from Liberia.

3.3 *Adjective*

The high frequency of copula absence for adjectives that Holm first found in the Jamaican data emerges even more dramatically in my reanalysis, as zero climbs from 66% (in his study) to 81% (in mine). The difference seems to lie in his inclusion of several tokens of *a* which I discounted — cases before *so* and *how*, perhaps, which I classed as adverbs and put in the “miscellaneous” category of Table 26.

The classical creole analysis of adjectives (Bickerton 1972:648, Holm 1984:295-6) is of course as a sub-type of stative verb, which from this point of view would no more require a copula than a stative verb (e.g., *know*) would. The justification for treating Adjective as a special category, however, is that it does occur with overt *be* copulas some of the time (18%, representing reanalysis), and that, as Bailey (1966:42) pointed out, adjectives differ from true verbs in several syntactic respects, including their co-occurrence with intensifiers (*so*, *thus*).

3.4 *Verb (+ed)*

Column 4 in Table 3 shows a separate analysis of stative Verb (+*ed*) predicates — forms like *engage* and *recommended* which many linguists classify with adjectives in their copula analyses as a matter of course. (I'm not sure

what Holm did with them; since most of these are passives, they might have been excluded.) Note that the statistics in the *Adj* and *V(+ed)* columns match very closely. If combined, the percentage of zeros for 'Adj/V+ed' would be 79%.⁶

3.5 *Verb+ing and Verb (+continuative)*

We come now to the heart of the differences between Holm's analysis and mine — the *V+ing* and continuative verb predicates that occur next in Table 3. As suggested earlier, I have long suspected that the 'Verb' statistics that Holm compared with the 'Verb+ing' of African American Vernacular English were an improper mixture of *V+ing* and *V+ing*. That this is so is clear from Table 5a, where I collapse these two categories in my data and reproduce his figures almost exactly. Collapsing the categories is improper, however, not only because they are syntactically incommensurate, but because their copula patterns are as different as chalk and cheese. *V+ing* occurs with \emptyset 86% of the time, while continuative verbs without an *-ing* suffix NEVER do so, taking instead the creole continuative markers *a* and *de* 86% of the time. Except for one case of *de V+ing* ("him de digin a di katin tri" 666) the two predicate types are almost in complementary distribution. The emergence of \emptyset *V+ing* represents a reanalysis of creole *de V*. The two predicates should be separated in a variation analysis, and only the *V+ing* statistics are properly comparable with *V+ing* in AAVE.

3.6 *gwain V and go V*

A similar argument applies to the future marker. Table 5b shows that Holm's low \emptyset figure for *gwain* — which has puzzled me for years — represents a conflation of tokens of *gwain V* (which occur categorically with \emptyset) and *go V* (which occurs almost as categorically, with *de*). The 35% \emptyset which we get for

Table 5a. *V+ing* comparison

	Holm		Rickford	
	<i>V+ing</i>	<i>V+ing</i> & <i>V+contin.</i>	<i>V+ing</i>	<i>V+contin.</i>
\emptyset	17%	17%	17%	17%
<i>a</i>	6%	9%	9%	9%
<i>de</i>	76%	72%	72%	72%
<i>be</i>	2%	2%	2%	2%

Table 5b. *gwain* comparison

	Holm		Rickford	
	<i>Gwain</i>	<i>Gwain V</i> & <i>Go V</i>	<i>Gwain V</i>	<i>Go V</i>
\emptyset	32%	35%	35%	35%
<i>a</i>	—	1%	1%	1%
<i>de</i>	(68%?)	64%	64%	64%
<i>be</i>	—	0%	0%	0%

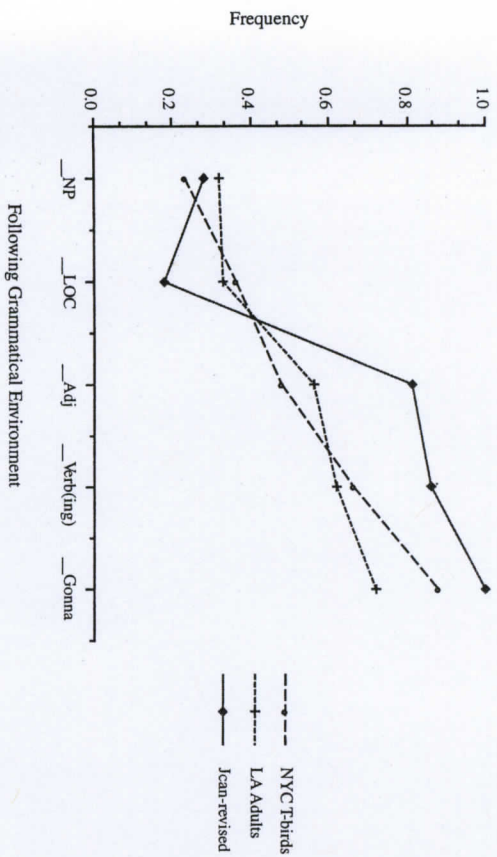


Figure 2. Copula Absence in 3 African-American Dialects (Jamaican revised)

this conflation in Table 5B comes entirely — as Table 3 shows — from tokens of *gwaïn*, the form equivalent to *gonna* and *gon* in representing a frozen reduction of *go+ing*. Interestingly enough, the *go* V cases behave almost exactly like the continuative verb cases insofar as they occur exclusively with *de* and *a*; they are of course just a special case of continuatives like *im de go wawk*, corresponding to “He is going to walk” in Standard English.

3.7 Copula absence by following syntactic environment reconsidered

If we now consider only the equivalent syntactic categories in AAVE and JC, their zero copula figures match even more closely, as shown in Figure 2 (JC percentages there based on Table 6). Furthermore, even the minor *NP/Loc* disparity in the comparison would be ironed out if, as suggested earlier, we compared their relative frequencies of *be*, rather than \emptyset , treating copula variation as an insertion rather than deletion process.

3.8 Variable rule analysis of the grammatical constraints on copula absence

I wish to turn briefly now to the results of a variable rule analysis of \emptyset in the first five subcategories of Table 3.⁷ Table 7 shows the results. The most

Copula Variability in JC and AAVE

Table 6. Copula variants by following grammatical environment in DeCamp's texts (my reanalysis, eliminating *V+con* and *go* V, and collapsing *Adj* and *V(+ed)*)

Variant	<i>NP</i> n = 68	<i>Locative</i> n = 40	<i>Adj</i> and <i>V(ed)</i> n = 82	<i>V+ing</i> n = 21	<i>Gwaïn</i> V n = 25
\emptyset	28%	18%	79%	86%	100%
a	18%	0%	1%	0%	0%
de	0%	65%	0%	5%	0%
be	54%	18%	18%	0%	0%

N = 236

Table 7. Varbrul run for JC copula absence, sharing significant factor groups

Input: 0.52	Personal Pro: .60	Other Pro: .23	NP: .70
SUBJECT:	NP: .23	Loc: .12	Adj: .75
FOLL. GR. ENV.:	NP: .23	Loc: .12	Adj: .75
TENSE:	Present tense: .61	Past tense: .39	V(+ed): .69
			V+ing: .79

significant factor group was FOLLOWING GRAMMATICAL ENVIRONMENT, with the factors following the order *V+ing* (most favorable to copula absence), *V(+ed)*, *Adj*, *Loc*, and *NP* (least favorable to copula absence). This at least establishes that the following grammatical hierarchy that emerged from the percentage figures in Table 3 is robust, and not the effect of other intersecting factors. The second selected factor is SUBJECT, with ‘other pronouns’ (forms like *dar*, *der*, *wat*, *wich* and *hu*) strongly disfavoring zero copula; one reason for this may be that *is* often occurs in contracted forms like *das*, although *dat iz* and *wat iz* do occur). For some reason a preceding NP is the most favorable environment for zero copula, unlike AAVE studies in which a personal pronoun is the most favorable.⁸ Note, however, that, in contrast with AAVE studies, ‘personal pronouns’ do not all end in a vowel, since *him*, *which*, and *it* occur also as subject pronouns in JC. The third selected factor is TENSE, with present tense favoring \emptyset over past, as we’d expect, but note past tense still allows a healthy amount of \emptyset (unlike AAVE).

Not selected by the regression analysis were person (whether the subject was first, second/plural, or third; AM never occurred in the data), and text (1-4). Coded, but not included in the variable run, were the effects of preceding and following phonological environment.

4. Summary and Conclusion

In a complete recoding and reanalysis of copula absence in De Camp's 1960 texts, I essentially replicate Holm's (1984) low zero copula percentages for a following Noun Phrase and Locative and his high zero copula percentages for a following adjective. In the *__Verb+ing* and *__gwan/gon(m)* V category, however, I find dramatically high zero copula percentages once syntactically non-equivalent forms are peeled off, and the resulting copula deletion hierarchy for Jamaican becomes dramatically more similar to that of AAVE, reinforcing arguments for the creole origins hypothesis.⁹ On the other hand, there is a reanalysis process that takes place for locatives and other forms, and the effects of Subject and Tense do NOT operate in JC in quite the same way they do in AAVE. Person is also irrelevant to copula absence in JC, but is of crucial importance for the corresponding variable in AAVE. We will clearly need to examine other data sets from Jamaica (see Rickford 1991 and Rickford, to appear, for a start), and we will need to follow the evolution from basilectal to upper mesolect in Jamaica in more detail than DeCamp's texts permit us to. But in the process of providing the crucial quantitative evidence that we need to understand variability in the JC and AAVE copula for its synchronic and diachronic significance, we cannot afford to neglect fundamental issues about how to define and count tokens of our variables, or we may see oases which turn out to be mirages, and we may miss mountains that are literally staring us in the face.

Notes

- * This paper is a revised version of one originally presented at NWAWE-19 in October 1990 at the University of Pennsylvania. It is a pleasure to acknowledge the assistance of Renee Blake and Angela E. Rickford. The paper was prepared while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, at Stanford, and the financial support provided by NSF Grants BNS-8700864 and BNS8913104 is gratefully acknowledged.
- 1 For the benefit of those who consult the published version of Rickford & Blake (1991), it should be noted that the order of pages 261 and 262 should be reversed.
- 2 In 1988, three undergraduate students of mine at Stanford — Jennifer Knobel, Diana Loo and Michelle Robinson — attempted, at my suggestion, a recoding and analysis of copula variability in DeCamp's data, but their useful preliminary work could not be incorporated in this paper for several reasons. One of the most important was that JC past tense and other tokens of the copula which were excluded in line with earlier analyses of

AAVE should not have been, since they were almost as susceptible to zero and other copula variants as present tense tokens were. The students' unfamiliarity with creole speech also led to a number of missed tokens and coding errors.

- 3 See Blake (1994) for a comprehensive review of the issue of 'don't count' cases in the analysis of copula absence in AAVE.

- 4 There were also 7 cases of actual or potential non-finite *be*, as in "yu wuda Ø supraiz," line 347, which I've excluded from the DC count because non-finites are not usually included in discussions of the copula anyway.

- 5 Holm makes this claim too in relation to *was*, but I don't think it applies to this form, which never, for instance, precedes (non-passive) verb stems. That is, one finds "im bin waan di tri gyal" (527), but not *"im waz waan di tri gyal," equivalent to "He (had) wanted the three girls."

- 6 Note incidentally that overt *-ed* tokens occur in only two cases, *contented*, *recommended*, both following denials and therefore syllabic, and that there's one token of *duon* ('*t*) for *wasn* ('*t*) in the data.

- 7 The last three subcategories all show categorical copula absence or non-absence (that is, Ø=100% or 0%) and would serve as knockout factors in the variable rule analysis, needing to be removed from the data pool before further analysis could continue.

- 8 This point is discussed in more detail in Rickford (to appear).

- 9 In a new data set examined in Rickford (1991) — drawing on the speech of Jack and Gertrude Harris (pseudonyms) of St Mary, Jamaica, recorded in 1991 — these results are essentially replicated, except that *__V+ing* shows about the same frequency of copula absence as *__Adj*. See Rickford (to appear) for these results. Quantitative analyses of copula absence in other Caribbean Anglophone creoles have appeared recently, e.g. Rickford (1992) on Barbadian and Winford (1992) on Trinidadian Creole, and these also show closer similarities with AAVE and boost the case for the creole hypothesis.

References

- Bailey, Beryl. 1966. *Jamaican Creole Syntax: A transformational approach*. Cambridge: Cambridge University Press.
- Baugh, John. 1979. "Linguistic style shifting in Black English". Ph.D. dissertation, University of Pennsylvania.
- _____. 1980. "A re-examination of the Black English copula". *Locating Language in Time and Space* ed. by William Labov, 83-106. New York: Academic Press.
- Bickerton, Derek. 1972. "On the nature of a creole continuum". *Language* 49, 640-69.
- Blake, Renee A. 1994. "Resolving the don't count cases in the quantitative analysis of the copula in African American Vernacular English". Paper presented at NWAWE23, Stanford University. To appear in *Language Variation and Change*.
- DeCamp, David. 1960. "Four Jamaican creole texts with introduction, phonemic transcriptions and glosses". *Jamaican Creole (Creole Language Studies 1)* ed. by Robert B. Le Page & David DeCamp, 128-179. London: Macmillan.
- Holm, John. 1976. "Copula variability on the Afro-American continuum". Conference

- Preprints, first annual meeting of the Society for Caribbean Linguistics, Turkeyen, Guyana. Compiled by George Cave, University of Guyana.
1984. "Variability of the copula in Black English and its creole kin". *American Speech* 59(4):301-309.
- Labov, William. 1969. "Contraction, deletion, and inherent variability of the English copula". *Language* 45:715-62
1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
1982. "Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor". *Language in Society* 11:165-201.
- Rickford, John R. 1991. "Variation in the Jamaican creole copula: New data and analysis". Paper presented at the Beryl Bailey Symposium at the Nineteenth Annual Meeting of the American Anthropology Association.
1992. To appear. "The development of AAVE". *African American English* ed. by Salikoko S. Mufwene, John R. Rickford, Guy Bailey & John Baugh. London: Routledge.
- Arneha Ball, Renee A. Blake, Raina Jackson, & Nomi Martin. 1991. "Rapping on the copula coffin: Theoretical and methodological issues in the variable analysis of the copula in African American Vernacular English". *Language Variation and Change* 3:103-132.
- & Renee A. Blake. 1990. "Copula contraction and absence in Barbadian English, Samana English and Vernacular Black English". *BLS* 16:257-268.
- Singler, John Victor. 1991. "Copula variation in Liberian Settler English and American Black English". *Verb Phrase Patterns in Black English and Creole* ed. by Walter F. Edwards and Donald Winford, 129-164. Detroit: Wayne State University Press.
- Winford, Donald. 1992. "Another look at the copula in Black English and Caribbean Creoles". *American Speech* 67:1:21-60

Contraction and Deletion in African American Vernacular English: Creole History and Relationship to Euro-American English

Ralph W. Fasold & Yoshiko Nakano
Georgetown University

1. Introduction

As Rickford et al. (1991) have put it, the contraction and deletion of *is* and *are* is a "showcase variable" in language variation analysis. This "showcase" status is very largely due to Labov's powerful analysis of contraction and deletion in the African American Vernacular English of New York City (Labov 1969, 1972). According to Labov's analysis, deleted forms of the copula and auxiliary were the result of contraction of those forms, followed by deletion of the remnants of contraction. His work suggested that the linguistic theory of the day, if expanded to allow for the careful, quantitative examination of actual speech data, would give convincing, orderly and intuitively satisfying results that could not be achieved with data from acceptability judgments alone. The results were the more impressive because they offered an account for several disparate observations.

- Why deletion of *are* is possible in white American English (Euro-American English), especially Southern Euro-American English, while deletion of *is* is common in AAVE but rare in Euro-American varieties (because in the relevant Euro-American dialects, *are* does not contain final [r] at the point of application of contraction, therefore contraction, which removes the initial vowel, removes the entire form. (The actual deletion rule, which is necessary to remove *is*, is a rule generally limited to AAVE.)

- Why the grammatical constraints, particularly the following grammatical constraints, on contraction and deletion taken together match the following environment constraints on contraction in a neighboring Euro-American