THE PHILADELPHIA STORY IN THE SPANISH CARIBBEAN

SHANA POPLACK AND DAVID SANKOFF
University of Ottawa Université de Montréal

THOUGH ESSENTIALLY MONOLINGUAL and Spanish-speaking, the Dominican Republic contains enclaves of native English speakers whose roots there go back to the early 1820s.¹ The historical record shows that several regions of the country were settled by some 6,000 American ex-slaves or their descendants who immigrated there through arrangements between the Haitian rulers of Santo Domingo and American church and philanthropic agencies (Commission of Inquiry 1871; Hoetink 1962; Rodríguez Demorizi 1973; Puig Ortiz 1978). Typhus soon decimated many of these settlements, but a number still exist.

In this article we describe research among residents of one such community in the peninsula of Samaná, which is largely separated from the rest of the country by a gulf of the same name. In 1981 and 1982 we tape-recorded conversations with nineteen people ranging in age from 71 to 103, who referred to themselves as “Americans” and spoke fluent English, some to the apparent total exclusion of Spanish, even in the third and fourth generations. Nearly all cited Philadelphia, New York, or New Jersey as the place of origin of their ancestors. Though there is ample documentation of sailings from these northern ports to Haiti, many of the passengers likely originated from various other parts of the United States as well.

Aside from the intrinsic interest of studying an enclave such as this one, the data collected there bear on two separate but interrelated issues: (1) the effects of over a century and a half of contact with Spanish on English, in one of the comparatively rare cases in which the latter represents the minority language, and (2) to the extent that the community had been isolated from, or resistant to, outside influence, analysis might provide some insight into the English spoken by black Americans early in the previous century. This in turn can provide us with a kind of evidence which may contribute to the longstanding debate on the genesis of American black English (ABE): Is its synchronic form due to gradual decreolization from a widespread creole similar to Caribbean creoles? Or were the earliest supraregional black varieties already rather close to white American dialects, so that ABE has actually been diverging away from standard American English (SAE)?²

For a detailed overview of the linguistic and sociological issues in-
volved in the creolist versus standard-origin positions on this debate, we refer the reader to Pfaff (1971), Wolfram (1974), Rickford (1977), Labov (1982) and Montgomery and Bailey (1986). Most relevant to the present paper are the contrasting tenets of the two positions that ABE and SAE were more divergent at earlier times (the creolist position) or of divergence comparable to (or even less than) that of modern times (the standard-origin position). In its extreme form, of course, neither position is tenable, especially with regard to the homogeneity of ABE in early times, but it would be helpful to establish the existence, in specific historical periods and geographical regions, of populations whose language can be characterized as more like documented creoles or more like contemporary white dialects.

If we can establish the extent of outside influences on Samaná English, and, in particular, influences from the external English-speaking world—since it is these which most threaten to invalidate our data as evidence about early ABE in the United States—we can assume that we are dealing with a variety of English as it was spoken in Samaná around the turn of the century (when our informants were acquiring it), or about 75 years after the Americans originally landed; and to the extent that it was free from influence from other varieties, it may represent a lineal descendant of ABE as it was spoken in Philadelphia, or in some wider region from which the immigrants were drawn, in 1824. Our analyses suggest that this variety was already highly decreolized, if indeed it ever represented a widespread, unified black American creole.

Evidence about (1) the social and geographical origins of the first input settlers, and (2) the linguistically relevant contacts they may have had with both Spanish and English-speaking outsiders, can be brought to bear on these suggestions. We begin by assessing the provenance of the input settlers. The Samaná tradition that most of the immigrants were recently escaped slaves is consistent with the possibility that at least some of them were newly arrived from the South. Newspaper accounts of the period (Niles Register, July 1825) reporting release of slaves from Virginia and North Carolina for direct exportation to Haiti lend further support to the suggestion that the first input settlers included, but were not limited to, northern blacks.

Even less is known of their social origins. A few reports of direct manumission of entire plantations to Haiti (ibid.) or to New York for immediate embarkation to Haiti (Maryland Gazette, Dec. 1824) indicate that both field and house slaves were among the original settlers. We have no way of knowing, however, the exact port of disembarkation of the immigrants, since specifics were not provided beyond the general destination
of "Haiti," nor is it known if any particular group of settlers better survived the widespread incidence of typhus.

The historical record is only a bit more revealing with respect to subsequent contact with other languages and/or varieties. Aside from the normal processes of internal evolution inevitable to all speech communities, there have been at least three potential external sources of linguistic change in Samaná English. The first and most evident is Spanish, the dominant and official language of the region for at least the last 110 years, as can be inferred from inhabitants' testimony (Commission of Inquiry 1871), but probably ever since the late 1840s after the replacement of the Haitian government by an independent Dominican authority. Another source might be the range of Haitian creoles and other French varieties to which the first settlers were exposed during the generation 1824–1845, creoles which survive in a variety of Haitian Creole the Samanese call *patois*, spoken by only a minority of our elderly informants. Third, and most problematic in the context of this research, is contact with individuals from English-speaking populations, Britain, the United States, Jamaica, and other Caribbean communities.

We begin by assessing the effects of contact with Spanish, first in terms of the massive language shift from English to Spanish which is today almost complete in Samaná, and then in terms of its implications for the structure of the English which is still spoken natively by the oldest generation. We will then evaluate the extent of contact with other varieties of English. (The influence of *patois* on Samaná English, which we feel to be detectible at this time only in a few lexical items, e.g., *zavoka* 'pear', will not be a focus here.) The main body of this paper will be devoted to the study of a single complex grammatical topic, copula contraction and deletion in Samaná English, and its pertinence for the historical questions which concern us.

Contact with Spanish. A number of outcomes are conceivable where speakers of the minority language (in this case English) are isolated and maintain restricted or no contacts with other native speakers:

1. Descendants of the original speakers of the minority language may show an increasing demographic shift so that it loses out to the majority language, Spanish (Language shift leading to language death).

2. The minority language may survive in a form heavily influenced by borrowings from the majority language (a mixed language).

3. Its domain of use could become progressively more restricted, possibly accompanied by simplification in linguistic structure (a sort of autonomous creolization).
4. It may undergo normal internal evolution, retaining some archaic features and innovating others, with little debt to or influence from other dialects of either Spanish or English.

5. If, on the other hand, it is in continuous contact with neighboring dialects of English, it may evolve in such a way as to parallel their syntactic structures. This is known as linguistic convergence, and in the case at hand may, depending upon the contact involved, result in either creolization (contact with Caribbean creoles) or decroelization (contact with more standard varieties of English).

Which of these tendencies predominate in the Samaná context? Although our informants are third and fourth generation residents of the Dominican Republic, they speak English almost exclusively. In fact, some claimed not to speak any Spanish, though later observations of their interactions with monolingual Spanish speakers indicate that this claim is somewhat of an exaggeration. Extensive bilingualism was documented by the American Commission of Inquiry as early as 1871. Nevertheless, as far as the oldest speakers are concerned, English can be said to be used in virtually all domains within the community, so that the results of restricted usage are not at issue. Their children (for the most part in their sixties) are bilingual, while their grandchildren speak mainly Spanish. Many say they have, with little result, forbidden their grandchildren to speak Spanish with them, hoping to maintain English a little longer. Most agree, then, that when they die, English will die with them:

My parents never spoke one Spanish word with me and I never spoke one Spanish word with my children. (MO)

I always tell them, "Why you all are English-speaking people, and you speak Spanish with your children? Speak the English with them!" You believe? But they don't. (ES)

With the tourists what's comin' in the town, the little Spaniards is breakin' their head to learn English and our'n is refusing it. (MS)

Thus language shift is a clear outcome, with the prospect of the ultimate loss of the particular variety of English characteristic of Samaná. Similar conclusions were drawn by Vigo, who administered a questionnaire to a random sample of residents of several communities in the Samaná region (Vigo 1986). This is why we have limited our investigation to the oldest (and last) generation of native speakers of Samaná English. We make no claims about the behavior of any of the younger generations, who in any event have incomplete or non-native acquisition histories of English.

Despite more than a century and a half of contact, Samaná English shows only superficial influence from Spanish. Relatively few established
loanwords recurred in the speech of our informants: Aside from terms designating Dominican institutions (e.g., *universidad* 'university'; *sanidad* 'health department'; *tribunal* 'court', and *ayuntamiento* 'town hall') which can be interpreted as proper nouns, there are a small number of grammatical elements, for example, the adverbs *ya* 'already' and *de repente* 'suddenly', and the conjunctions *ni* 'nor' and *o sino* 'or else'. Other loanwords include the ubiquitous introducer *bueno* 'well', and the lexical items *turismo(s)* 'tourism/tourists', *carretera* 'road', and *chinos* 'oranges'. All of these were pronounced with English phonology, as were most Dominican place names mentioned (e.g., *Porto Plat* < *Puerto Plata*, *Fletch* < *La Flecha*, *Ol' Hat* < *Hato Viejo*). There is an occasional tendency to form calques on Spanish constructions (e.g., *gain money* < *ganar dinero*) as well as some phonological transfer which, however, varies according to the informant. This fairly well resumés the extent of Spanish influence to be found in our corpus; there has been no massive influx of either lexical items or grammatical constructions. The mixed (Spanish-English) language scenario is not one of the outcomes of language contact in Samaná.

**Contact with Dialects of English.** The Wesleyan Methodist Church and, later, the African Methodist Episcopal Church appear to have been in part responsible for maintaining the distinctiveness of the American community and its language. Input from the outside, however, was extremely limited, especially in the last century. The archives of the Wesleyan Methodist Missionary Society reveal that there was a British missionary in Samaná during much, if not most, of the period from 1838 through the 1930s (Vigo, personal communication). His example would surely constitute a standardizing influence on Samaná English, but it is unclear how many of the settlers might have had significant exposure to such influence. Correspondence of several of the early ministers with the society includes complaints that few of their American parishioners could read or write (ibid.). This suggests that the influence of the teachers/ministers did not extend beyond the small proportion of individuals attending school for more than a few years (for example, we did note occasional lexical items of Jamaican provenance in the speech of a local preacher whose elementary schooling had been with a Jamaican). We do not, however, believe this to have had more than an indirect and attenuated effect on the speech patterns of the population as a whole. Indeed, there is some evidence of internal sociolinguistic differentiation of the community. Hoetink (1962) states "There clearly exists an elite, living in the little town itself. . . . Their English, somewhat archaic perhaps and
with an elaborate use of biblical parables, is much better understandable than that of the isolated farming people in the surrounding area." We too will show evidence of sociolinguistic stratification below.

There were also intermittent visits from white missionaries from the British possession of Turks Island about 500 miles north; at least one or two local preachers attended church synods in Haiti, and one man was actually sent to Philadelphia to be trained at a Methodist college. As well, there appears to have been a number of American businessmen, English missionaries, and Jamaican preachers or school teachers in residence at various times, and some immigration, trade, and intermarriage involving St. Thomas and, especially, Turks Island. The English spoken by these latter contacts would have represented whatever points on the creole continuum existed on those islands at that time. There is also evidence of interaction with other American settlements at Puerto Plata, Santo Domingo, and elsewhere, although these were apparently Hispanicized relatively early on.

These contacts, however, were spread out over several generations and at no time were there more than a few foreign native speakers of English present in the community. Moreover, with the exception of those directly involved in education and/or the ministry, none were likely to have had a significant impact on the vernacular spoken in the scattered network of agricultural hamlets which constituted the American settlement, let alone to have exercised the strong, sustained, decroolizing influence which might be adduced to account for the type of English now spoken in Samaná.

Indeed, if Samaná English has (at least until the present) so successfully resisted convergence with Spanish, the language which has indisputably dominated and surrounded the English-speaking enclaves in the Dominican Republic for over a century, how likely is it that the much more limited contacts with the external English-speaking world have transformed this variety, say from some basilectal or mesolectal creole, to something approaching modern ABE?

On the contrary, the geographical remoteness and relative isolation of the peninsula of Samaná during the period which interests us (before the First World War, when our informants had already acquired or were acquiring their present-day English) weakens the hypothesis of externally motivated linguistic change (i.e., convergence or decroolization) considerably. Indeed, these are precisely the characteristics associated with the concept of linguistic enclave or isolated area which are widely accepted in historical linguistics and dialect geography to retain conservative features. Thus we conclude that, unless there are key facts of
which we are unaware, it is improbable that outside English dialects had any major effects on Samaná English. We will see below that the purely internal linguistic evidence we bring to bear is fully consistent with this claim.

A cursory examination of the existence in Samaná English of a series of features generally considered to be representative of ABE and West Indian creoles (as abstracted from inventories provided by Stewart 1967, 1968, 1970; Brewer 1973; Bickerton 1975; Rickford 1977; Lourie 1978, and others) bolsters our analysis, revealing the following:

1. Features which are specifically characteristic of English-based creoles, particularly in the areas of morphology and syntax, and most strikingly in the tense and aspect system (Tagliamonte and Poplack 1986), appear rare or nonexistent in Samaná English.

2. By the same token, although Samaná English shares with ABE many differences from standard English, these are generally no more prominent or more frequent in Samaná than in modern ABE. This is not what we would expect if ABE were a decroлизed descendant of an earlier variety and Samaná an offshoot of the same variety but relatively unaffected by decroлизing pressures. The remarkable fact that Samaná English has remained relatively free of influences from outside sources makes it an ideal test case for the synchronic investigation of features associated with at least one variety of early ABE.

We approach this problem with a comparative study of perhaps the best-documented feature in ABE dialects, contraction and deletion of the copula. If Samaná English and, ipso facto, older stages of ABE were creolelike, we would expect Samaná English to evidence patterns of copula deletion quantitatively more advanced than those of present-day ABE, and more similar to copular configurations of English-based West Indian creoles.

**Contraction and deletion of the copula.** By copula we refer to the present tense of the verb to be before predicate nouns, adjectives, locatives, and comitative phrases. It is well known that in ABE (see Fasold 1969, Pfaff 1971, Labov 1972, Brewer 1973, Butters 1973, Wolfram 1974, Baugh 1980) the copula is often absent in these syntactic environments, as well as when be is used as an auxiliary. The following examples of contexts in which the copula may be deleted show that this behavior is also characteristic of Samaná English:

1. ____/Noun Phrase:
   This young man, his father and his mother $\emptyset$ good English-speaking people, and he don't speak it correct. (MO/371)
2. ______/Predicate Adjective:
   When they draw your profile with your workin' clothes on, it's better
   than when you Ø dress'. (MO/110)

3. ______/Locative:
   If anybody Ø in the way, well, they'll mash him up. (IC/275)

4. ______/Verb + ing:
   He Ø j'ining the army. (SJ/423)

5. ______/gonna:
   She don't ax me what she Ø gon' cook. (MS/82).

Such examples of copula absence have led many observers (e.g., Stewart 1967, 1968) to conclude that there is no present copula or auxiliary be in ABE. This would seem reasonable in view of the fact that many languages do not have a copula, particularly some to which ABE may be related. However, in Labov's (1972) analysis of the natural speech of black peer groups in Harlem, he showed that although the copula and auxiliary be were frequently absent in environments such as the ones listed above, infinitival, emphatic, or exposed copulas were never deleted. These may also be exemplified with data from Samaná, as in the following examples:

1. In questions:
   You see where that coconut tree is down there? (C/3)

2. In comparative constructions after ellipsis:
   English ain't so easy to learn like Spanish is. (MS/2)

3. In embedded questions after wh:
   I suppose you don't know what that is. (ES/249)

4. When the copula is infinitival or follows a modal:
   Well, when we die, the oldest ones, the English will be
   scarce around here. (SJ/122)

5. Under emphasis:
   The water I think is going to break a little. (ST/11)

Labov's quantitative analysis of his data showed a strict relation between standard English patterns of copula contraction and deletion in ABE: wherever standard English can contract the copula, ABE can delete it; and conversely, wherever standard English cannot contract, ABE cannot delete. He found that contraction and deletion of the copula form is were constrained by both phonological and grammatical factors. One important constraint was the nature of the preceding noun phrase—if this noun phrase was a subject pronoun (he or she) both contraction and deletion of the copula were likely. He also examined the grammatical categories following the copula, and ranked them according to how likely contraction and deletion were to take place in these environments, as schematized in figure 1, which illustrates the frequency of contraction and deletion of the copula among the Cobras in Harlem.
This ranking parallels the ranking with which copula contraction occurs in standard English, suggesting that copula absence or deletion is simply an extension of the standard English contraction process. This ran counter to the creole-origin hypothesis whereby ABE, like many creoles, would have no underlying copula in some environments, or a copula system otherwise very different from that of standard English. According to this hypothesis, the variability we observe in ABE today would be due to (irregular) copula insertion in the process of decreolization. The similarity of contraction and deletion patterns would seem incompatible with the decreolization explanation since the parallel ranking of following grammatical influences on contraction and deletion in ABE and standard English would remain unexplained, and far too coincidental.

Rickford (1977) suggested that this evidence might not be too damaging to the creole-origin hypothesis in speculating that the parallel ranking found by Labov simply reflects universal preferences for copula presence in nominal-versus-attributive-versus-verbal contexts. The parallelism would thus be irrelevant to the debate over historical origins. However, data provided by Holm (1984), a paper first read in 1976, reopened the question by showing that the ranking is not universal, and that at least two creoles have a dramatically different pattern of copula use from that published by Labov for ABE.5

As illustrated in figure 2, Jamaican Creole and Gullah have extremely high rates of copula absence before an adjective, as opposed to the low
rates calculated for this context in the Harlem study (figure 1). Indeed, Holm distinguished between following locative and adjective contexts, and predicted for ABE that, were adjective and locative treated separately, the adjective would appear much more often without the copula than would the locative, as in the creole data. Baugh (1980) took up this suggestion, and separated the adjective and locative categories in a more complete set of the data on one of the adolescent groups (the Cobras) studied by Labov. He did in fact find a clear difference between locative and adjective contexts, but not the one predicted by Holm. In fact, he found that adjectives appear with the copula far more than do locatives, and almost as much as do noun phrases. This is the opposite of what is found in creoles and can be considered evidence against a close relationship between creoles and ABE. Under the hypothesis that the contraction rule operates prior to deletion, however, Baugh found that, whereas adjectives disfavor contraction of the copula, they do favor the deletion of any contracted ones. He attributed the contraction results to standard English influences, while the deletion rule was inferred to contain vestiges of creole origins. The creole syntactic rule of pre-adjectival copula absence is then considered to have been replaced by a non-creole pattern of categorical underlying copula susceptible to a two-stage phonological reduction process, only the second stage of which reflects
creole patterns. This multistep reasoning may prove valid, but given the very small number of contracted or deleted pre-adjectival tokens (less than ten?) from which the phonological effect could have been calculated, it must be thoroughly tested and replicated before it can be considered to outweigh the simple fact of high rate of copula presence in this environment, which is counter to the creole pattern. Do Baugh's results hold for copular forms other than is? If high deletion rates in pre-adjectival position are a vestige of creole syntax, this should be true of are and am as well (at least in comparison with locative contexts). Does Baugh's analysis hold for larger data sets or in other ABE communities? In a replicate study in Los Angeles, Baugh (1979) did find a tendency for pre-adjectival is to resist contraction and to be susceptible to deletion, but this was not as clear as in the Harlem data. Furthermore, it did not hold for are, which resists both contraction and deletion in pre-adjectival position.

It is in the light of these questions that we examine our Samaná data. If the creole-origin hypothesis is correct, then Samaná English, insofar as it represents a variety of ABE as spoken a century and a half ago (and we have presented here historical evidence that this is likely to be the case), should show far more differentiation from standard English than the presumably decreolized ABE spoken in Harlem today (given the strong form of the decreolization argument which assumes that the farther back in time one goes, the greater the differences between ABE and the standard). So in the case at hand, the Samaná data should show far more copula deletion than data on ABE. Similarly, if Samaná English were decreolizing, or if it were converging with surrounding West Indian creoles, a possibility we have not yet discounted despite indications to the contrary, we would expect to find a hierarchy of constraints similar to those attested for West Indian creoles (in particular, that a predicate adjective favours copula absence).

**Methodology.** We extracted nearly 500 instances of the copula (including zero realizations) from lengthy tape-recorded interviews we conducted in 1981 with eight of our elderly informants, all of whom are native speakers of English. We would have expected our recordings to tend toward the formal side of the stylistic range, given the differences between the interviewers and the informants in age, race, and general sociocultural background. We thus carried out standard sociolinguistic checks for detecting more vernacular styles, using the internal stylistic differentiation of the interviews, paying attention to jokes and other more spontaneous discourse modes, asides to other household mem-
bers, and background conversations among older English speakers not involving the interviewers. We were particularly concerned about the possible existence of a more basilectal variety. However, the only dramatic shifts in speech mode we could find, and this pattern was confirmed in our later fieldwork in 1982, were occasional switches to patois or Spanish to accommodate to non-English-speaking interlocutors.

**Coding Procedures.** For comparability in coding the data, we tried to use the same coding system as in previous work, but some changes seemed desirable.

For example, while I’m occurs over ninety-nine percent of the time in Labov’s data, making statistical analysis superfluous, here it occurs only eighty percent with ten percent full forms and ten percent deleted forms. Thus, we included am and its contracted and deleted variants in our analysis.

Similarly, Labov did not include data on are-contraction and deletion in his statistical analysis. He interpreted the almost categorical absence of are as resulting from vocalization of final r followed by loss (contraction) of the vowel of the copula. Wolfram (1974), however, has argued convincingly for a reduction process parallel to that for is, with contraction followed by deletion of the r. Our data, with a non-negligible number of full and contracted forms, is more consistent with Wolfram’s analysis, and so we include are and its variants in the statistical analysis.

In some dialects of black English, lack of person/number agreement between subject and copula is amply attested so that in, for example, You running it is not clear whether are or is has been deleted. But in Samaná as with modern ABE (Labov 1972, fn. 29, though not Labov and Harris 1986), there is only a low rate of person/number disagreement. Thus, even though we analyzed statistically all instances of is, am, and are together, we could take account of differential reduction of these forms by distinguishing between subjects on the basis of person and number.

Labov treated is’s, tha’s and wha’s as the outcome of the obligatory assimilation of t in it’s, that’s and what’s to the s, followed by the obligatory deletion of the copular s since it now would follow another s. Hence, he did not need to include these cases in his statistical analysis. Since we had a number of tokens like That is my daughter and That Ø my daughter we did include tokens with that, it, and what subjects. In coding these we assumed, contrary to Labov’s (1972) suggestion, that the final t-assimilation follows, not precedes, the copula-deletion rule; otherwise the That Ø sequences could not be realized.

With Baugh, we coded for both preceding and following phonological environment, and we distinguished noun phrases beginning with deter-
miners from those without in the following grammatical category. We also distinguished participles like finished and done from other predicate adjectives. Predicates beginning with a wh-form were coded separately, and we included a miscellaneous category containing largely adverbials and demonstratives.

**Data Tabulation.** Table 1 shows the overall contraction and deletion rates for the speakers in our sample. Our sample of speakers is of course too small to allow any definite conclusions, but it does seem to confirm our earlier observations that there is some internal differentiation in the speech community between those who have had more than a few years of schooling and/or lifelong association with church leadership, versus those who did not, the latter deleting copulas more than the former. Note, however, that deletion is fairly low across the board.

We now assess the percentages of copula presence in different linguistic environments. Table 2 shows the effects of different kinds of subject on the process of copula reduction. Concentrating on the he/she and full NP figures, we can compare the behavior of the Samaná speakers with information on other dialects of ABE, as in table 3.

The only data in table 3 that clearly show less deletion than the Samaná materials come from more formal speech styles or middle class speakers of northeastern ABE. Note, however, that the southern black folk speakers recently studied by Bailey and Maynor (1985) show similar rates of deletion. From this perspective, Samaná speech may be characterized as rather more similar to standard English with respect to copula deletion. This quantitative evidence runs counter to the impressions of DeBose (1984) that Samaná English is closer to the basilectal pole of the modern black English continuum. Bolstering our conclusion, we note from table 2 the non-negligible frequency of full forms of am, are, and {that, it, what} + is. These do not appear in such numbers in varieties in which copula reduction is more advanced. The fact that here, where, and

### Table 1

<table>
<thead>
<tr>
<th>Speaker</th>
<th>1*</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5*</th>
<th>6</th>
<th>7*</th>
<th>8*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>.21</td>
<td>.11</td>
<td>.19</td>
<td>.62</td>
<td>.20</td>
<td>.35</td>
<td>.20</td>
<td>.40</td>
</tr>
<tr>
<td>Contracted</td>
<td>.62</td>
<td>.62</td>
<td>.24</td>
<td>.23</td>
<td>.73</td>
<td>.37</td>
<td>.61</td>
<td>.60</td>
</tr>
<tr>
<td>Deleted</td>
<td>.17</td>
<td>.27</td>
<td>.42</td>
<td>.15</td>
<td>.07</td>
<td>.28</td>
<td>.18</td>
<td>0</td>
</tr>
<tr>
<td>N</td>
<td>107</td>
<td>100</td>
<td>26</td>
<td>13</td>
<td>85</td>
<td>65</td>
<td>93</td>
<td>5</td>
</tr>
</tbody>
</table>

* = Church/school identification
Table 2
Proportion of Full, Contracted, and Deleted Copula in Samaná with Different Pronominal and Full NP Subjects

<table>
<thead>
<tr>
<th></th>
<th>I am</th>
<th>he is</th>
<th>she is</th>
<th>it is</th>
<th>what is</th>
<th>that is</th>
<th>there is</th>
<th>here is</th>
<th>where is</th>
<th>are</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>.10</td>
<td>.02</td>
<td>0</td>
<td>.11</td>
<td>.07</td>
<td>.12</td>
<td>.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contracted</td>
<td>.80</td>
<td>.82</td>
<td>.74</td>
<td>.85</td>
<td>.71</td>
<td>.85</td>
<td>.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deleted</td>
<td>.10</td>
<td>.16</td>
<td>.26</td>
<td>.05</td>
<td>.21</td>
<td>.05</td>
<td>.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>80</td>
<td>51</td>
<td>34</td>
<td>66</td>
<td>14</td>
<td>82</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) (b) (c) (d)

<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th>we are</th>
<th>you are</th>
<th>they are</th>
<th>these are</th>
<th>those are</th>
<th>is</th>
<th>them</th>
<th>is</th>
<th>are</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>.76</td>
<td>0</td>
<td>.23</td>
<td>.26</td>
<td>.83</td>
<td>.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contracted</td>
<td>.09</td>
<td>.10</td>
<td>.10</td>
<td>.16</td>
<td>0</td>
<td>.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deleted</td>
<td>.15</td>
<td>.90</td>
<td>.67</td>
<td>.58</td>
<td>.17</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>74</td>
<td>10</td>
<td>30</td>
<td>19</td>
<td>12</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(e) (f) (g)

there can be associated with all copula forms may explain their pattern intermediate between the singular and plural pronoun results.

These figures basically confirm Labov's results as concerns the difference between full NPs which inhibit is-reduction, and certain pronouns which favor it. They also provide a quantitative generalization of the more categorical relationships he posited between other pronouns and specific stages in the reduction of the associated form of the copula,

Table 3
Proportion of Deleted is in Various Dialects of ABE and in Samaná English

<table>
<thead>
<tr>
<th>Group Studied</th>
<th>Proportion of Deleted is</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working class Detroit adults (Wolfram 1969)</td>
<td>.25</td>
</tr>
<tr>
<td>Harlem adults: informal speech (Labov 1972)</td>
<td>.14</td>
</tr>
<tr>
<td>Samaná</td>
<td>.15</td>
</tr>
<tr>
<td>Harlem adults: formal speech (Labov 1972)</td>
<td>.08</td>
</tr>
<tr>
<td>Middle class Detroit adults (Wolfram 1969)</td>
<td>.04</td>
</tr>
<tr>
<td>Lower class Texas adults (Bailey and Maynor 1985)</td>
<td>.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group Studied</th>
<th>Proportion of Deleted is</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working class Detroit adults (Wolfram 1969)</td>
<td>.25</td>
</tr>
<tr>
<td>Harlem adults: informal speech (Labov 1972)</td>
<td>.14</td>
</tr>
<tr>
<td>Samaná</td>
<td>.15</td>
</tr>
<tr>
<td>Harlem adults: formal speech (Labov 1972)</td>
<td>.08</td>
</tr>
<tr>
<td>Middle class Detroit adults (Wolfram 1969)</td>
<td>.04</td>
</tr>
<tr>
<td>Lower class Texas adults (Bailey and Maynor 1985)</td>
<td>.18</td>
</tr>
</tbody>
</table>
exemplified by *I am / I'm, that is / tha's, and you are / you.* The variation in our data appears consistent with the hypothesis that NP/pronoun and singular/plural are the subject distinctions pertinent to the reduction of the copula, though *am, is,* and *are* may have different (lexical) tendencies to contract, and phonological processes may have a differential effect on the deletion of the final consonantal segments *m, s,* and *t.*

For the statistical analysis in the next section, we will not retain the distinctions between all possible subjects; rather we combine the data on the similar kinds of subjects within table 2(b) together, and likewise with 2(c), 2(f) and 2(g).

Table 4 shows the data classified by following grammatical category.

Perhaps the most striking aspect of these figures is that the ranking of following NP, adjectives and locatives, *V + ing,* and *gonna* parallels Labov's original results. (Compare figure 3 with figure 1.) Note also that locatives, adjectives and participles all share virtually identical patterns. In particular, the so-called creole order, with high deletability of the

<table>
<thead>
<tr>
<th>wh-clause</th>
<th>gonna</th>
<th>V + ing</th>
<th>Loc</th>
<th>Part</th>
<th>Adj</th>
<th>DET + NP</th>
<th>NP</th>
<th>Misc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>0</td>
<td>.02</td>
<td>.16</td>
<td>.29</td>
<td>.21</td>
<td>.27</td>
<td>.26</td>
<td>.34</td>
</tr>
<tr>
<td>Contracted</td>
<td>.41</td>
<td>.56</td>
<td>.52</td>
<td>.62</td>
<td>.64</td>
<td>.58</td>
<td>.64</td>
<td>.52</td>
</tr>
<tr>
<td>Deleted</td>
<td>.59</td>
<td>.42</td>
<td>.32</td>
<td>.18</td>
<td>.14</td>
<td>.15</td>
<td>.10</td>
<td>.14</td>
</tr>
<tr>
<td>N</td>
<td>17</td>
<td>41</td>
<td>73</td>
<td>66</td>
<td>14</td>
<td>55</td>
<td>97</td>
<td>71</td>
</tr>
</tbody>
</table>

**Figure 3**
Frequency of Contraction and Deletion of the Copula
According to Following Grammatical Environment in Samaná English
copula before predicate adjective, certainly does not pertain here, with only fifteen percent deletion in this environment. (Compare with figure 2).

Because the finer categorization of the following grammatical category adopted from Baugh does not reveal any new quantitative distinctions in table 4, for the regression analysis in the next section we will recombine $\text{DET.} + \text{NP}$ and $\text{MISC.}$ with $\text{NP}$, and participles with adjectives.

**Multiple regression analysis.** Tabulations such as tables 1, 2, and 4 are very informative, but for a maximum of linguistic interpretability and statistical rigor, further manipulation of the data is necessary. First, the totals of contracted and deleted forms must be contrasted with the unreduced forms to study the contraction process which we hypothesize, with Labov, Wolfram, Baugh and others, to be ordered before deletion. Then the deleted forms must be compared to the contracted ones to characterize the deletion process. In each case a multiple regression analysis is used to estimate the true effects of social, phonological, and syntactic features on the reduction process, and to remove any artifacts of poor data distribution, correlated factors, or statistical fluctuations.

VARBRUL 2S, the computer program used to carry out this regression (Sankoff 1979), is based on the principle of maximum likelihood to determine which environmental factors have a significant effect on contraction and/or deletion when all are considered simultaneously, and to estimate the magnitude of the individual factor effects. Table 5 presents the results of the multiple regression analysis of copula contraction.9

The higher numbers favor contraction; lower ones disfavor it.

Some of the interesting results of this analysis are that neither following phonological environment nor speaker characteristics have a signifi-

<table>
<thead>
<tr>
<th>Subject</th>
<th>Preceding Phonological Segment</th>
<th>Following Grammatical Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$</td>
<td>.45</td>
<td>consonant</td>
</tr>
<tr>
<td>she, he</td>
<td>.93</td>
<td>vowel</td>
</tr>
<tr>
<td>it, what, that</td>
<td>.85</td>
<td>$\text{gona}$</td>
</tr>
<tr>
<td>here, there, where</td>
<td>.74</td>
<td>$\text{V + ing}$</td>
</tr>
<tr>
<td>$\text{NP}$</td>
<td>.08</td>
<td>Loc</td>
</tr>
<tr>
<td>$\text{we, you, they}$</td>
<td>.32</td>
<td>Adj</td>
</tr>
<tr>
<td>$\text{those, them, these, this}$</td>
<td>.13</td>
<td>$\text{NP, wh-clause}$</td>
</tr>
</tbody>
</table>

*Factors not selected: speaker, following phonological segment*

Corrected mean: .89
cant effect on contraction (as neither factor was retained in the analysis by the multiple regression procedure.). The largest effect is exerted by the nature of the subject. A full NP retards contraction (with an effect of only .08) while subject pronouns promote it, with *he/she* (.93) and *it/what/that* (.85) showing the largest effects. We infer that *is* contracts more than *am*, which in turn contracts more than *are*. The preceding phonological segment also has an effect, with a preceding vowel favoring contraction. This can be explained as a movement toward a preferred CV syllable structure (Labov 1972, Pfaff 1971)—if the segment preceding *is, am, or are* is a vowel, the copular vowel drops to avoid a VV sequence. The contributions of following grammatical category agree with Labov’s analysis, *gonna* favoring contraction, followed by *V + ing*, locatives, and adjectives, with NPs the most resistant.

In the deletion rule (table 6) it is the following phonological environment which has a significant effect, interestingly enough, and not the preceding one. The direction of the effect is again explicable in terms of favoring CV sequences.

In the subject category we note the almost complete retention of contracted *am*, as well as *is* after *that, it, and what*. On the other hand, a high susceptibility of the contracted form to be deleted can be noted for NP and plural pronominal subjects. While the discord of the NP results with previous accounts may be due to insufficient data (seven cases of

| Table 6 |
| Contribution of Factors Selected as Significant by the Multiple Regression Analysis of Copula Deletion |

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Following Phonological Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>Consonant .64</td>
</tr>
<tr>
<td>2*</td>
<td>Vowel .36</td>
</tr>
<tr>
<td>3 and 4</td>
<td>Following Grammatical Category</td>
</tr>
<tr>
<td>5*</td>
<td><em>wh-clause</em> .95</td>
</tr>
<tr>
<td>6*</td>
<td><em>gonna</em> .59</td>
</tr>
<tr>
<td>7* and 8*</td>
<td><em>V + ing</em> .46</td>
</tr>
<tr>
<td>Subject</td>
<td>Loc .23</td>
</tr>
<tr>
<td>I</td>
<td>Adj .19</td>
</tr>
<tr>
<td>he, she</td>
<td>NP .41</td>
</tr>
<tr>
<td>it, what, that</td>
<td>Factors Not Selected</td>
</tr>
<tr>
<td>NP</td>
<td>Preceding phonological segment</td>
</tr>
<tr>
<td>we, you, they</td>
<td>Corrected mean: .54</td>
</tr>
<tr>
<td>here, there, where</td>
<td>* = church/school identification</td>
</tr>
<tr>
<td>them, those, these, this</td>
<td>.53</td>
</tr>
</tbody>
</table>

* = church/school identification
deletion; four of contraction), or to the previously mentioned inclusion of plural NP's, they are not completely implausible. Together with the resistance to contraction evident in table 5, they simply indicate that a contracted copula attached to a full NP is not a favored form. This is quite understandable for certain NP's, especially those with the head noun in nonfinal position, and further studies of the topic with additional data would evaluate this hypothesis. If this (or any NP/pronoun effect) is confirmed, it is clearly a syntactic effect, but the differential behavior of singular and plural copula is due to phonological differences and/or quantitative lexical tendencies.

The interaction between particular pronoun and the form of the copula are individual relationships that may be in the process of taking on fixed forms: I'm, is, tha's, they Ø. We cannot with confidence separate out the effect of number and person from form of the pronoun. Thus the statistical results are determined by the fact that I'm, tha's, they Ø, etc., are favored configurations, though perhaps not as categorical as reported for ABE.

The following grammatical category behaves again as in almost all the ABE studies. Though following NP's do not disfavor deletion as much as adjectives and locatives, the verbal forms are the most amenable to deletion. In terms of these factors, then, contraction is virtually parallel to deletion, confirming once again that, with respect to the system of copula predicates, deletion is an extension and generalization of contraction.

The speaker differences we saw in the percentages in table 1 are reflected very clearly in the deletion analysis, confirming, as expected, that it is the process of deletion which has social significance in the community, in contrast with contraction.

Finally, it should be stressed again that, in the ranking of factors for deletion, we have predicate adjectives, NPs, and locatives as least-favorable environments: in fact, adjectives have the lowest value, the opposite of the creole order.

Comparison with other communities. Two additional data sets which have been examined recently are pertinent to our discussion. One is the data on copula usage of ex-slaves analyzed by Bailey from the WPA tape recordings (Bailey 1985). The recordings were made in the 1930s and 1940s with former slaves born between 1844 and 1864, about a generation after the first emigrants to Samaná. All were from the South and had limited geographical and social mobility. Bailey's data, which we depict in figure 4, show that copula deletion rates for the ex-slaves are comparable to the modern speech of southern black rural adults and children he also studied, with pre-NP position the least favorable, fol-
followed by locatives and adjectives, which both allow somewhat more deletion, and then by pre-gonna and the environment before V + ing, which are by far the most favorable contexts. Adjectives do appear to promote deletion somewhat more than locatives in the ex-slave data, but it is striking that the ranking of constraints still resembles the Samaná figures far more than the ones Holm found for the creoles.

It is also instructive to examine another data set on the English spoken by Liberian Settlers (Singler 1986a), descendants of a group of ex-slaves who began emigrating from the U.S. to Liberia at approximately the same time (1822) as the ancestors of our Samaná informants. The Liberian experience contrasts with the history of the Samaná community in that the former speakers presumably were (and still are) in contact with substrate languages, that is, indigenous African languages and, especially, the regional English-based pidgin, which antedated the arrival of the Settlers. It would not be surprising, then, to find evidence of external influence on Settler English. Indeed, Singler (1986b) shows that it has a creole-like system of tense and aspect.

Now, the basilectal and mesolectal levels of non-Settler pidgin speakers show high rates of copula deletion, higher in pre-adjectival position than preceding NPs or locatives, as in the creoles studied by Holm (Singler 1986a). Figure 5 summarizes the rates and conditioning of copula deletion among the Settlers. Comparison with the Harlem and Samaná
results shows that both contraction and deletion are further advanced for the Settlers, but the ranking of constraints is identical to that of the latter two varieties.\textsuperscript{11} There is no trace of the creole pattern of pre-adjectival copula favoring deletion, characteristic of the Liberian pidgins. Thus, we may infer that the Settler ranking is a vestige of an early ABE copula system.

**Conclusions.** The grammar of Samaná English has been remarkably resistant to Spanish influence, considering the sociodemographic character of the region. We have also argued here that it has not been altered by contact with external English varieties, except with respect to some minor features. There have undoubtedly been some purely internal developments and innovations, but it seems most likely that this variety of English is a lineal descendant of the language as spoken by many blacks in the United States in the early 1800s. If we are right, then that language was no more creolized than modern ABE, and, at least insofar as its copula usage is concerned, it bore no more resemblance to English-based West Indian creoles than modern ABE, and indeed, less. We reached this conclusion by comparing rates and conditioning of copula usage among Samaná and quantitatively studied adult ABE speakers in Harlem, Detroit, and rural Texas, and finding not only overall rates but, significantly, the constraint ranking, similar to those attested for those varieties, while quite different from the few creoles which have been studied quantitatively.\textsuperscript{12}
Of course, even proponents of the creole-origin position would not deny that there existed highly decreolized versions of ABE at that early period, especially among house slaves and/or northern blacks. As has been documented, however, and in contrast to Samaná tradition, it is quite likely that a good proportion of the emigrants to the Dominican Republic from northern ports were not local residents but originated in the South, and were representative of various classes of ex-slaves.

Added to this is the fact that present-day rural black communities in the South and, especially, the ex-slaves born in the middle of the nineteenth century, have deletion rates and grammatical-constraint rankings virtually identical to those of the Samaná speakers. Such similarities can hardly be coincidental.

NOTES

1. Earlier versions of this paper were read at the Nwave X conference in Philadelphia (1981) and at the seventh Simposio sobre la dialectologia del caribe hispanico in San Juan, PR (1982). Preliminary versions of the results appeared in 1983 as Report 1161 of the Centre de recherches mathematiques, and as Poplack and Sankoff (1984). We are grateful to Moses Shepherd and family for guidance and help with the fieldwork and our anonymous informants for their cooperation and hospitality. We thank Sali Tagliamonte for her collaboration on the Samaná project, and the following colleagues for comments and suggestions: John Baugh, Ronald Butters, Ralph Fasold, John Holm, William Labov, Salikoko Mufwene, John Rickford, John Singler, and William Stewart. We are indebted to José Vigo for sharing with us his extensive knowledge of Samaná and its history.

2. Throughout this paper we will use the terms standard English and ABE as if they referred to well-defined, homogeneous entities, which of course, they are not. There are, however, clear systematic contrasts in the United States between most modern varieties of vernacular black English and varieties, historical and modern, which may loosely be characterized as standard, and it is these contrasts which we invoke when we use this terminology. Whenever it is pertinent to specify historical stages or regional varieties, this will be made explicit.

3. Such evidence is unfortunately rather spotty, the most copious historical documentation of the settlement of Samaná being that accumulated by José Vigo, who kindly provided us with the newspaper accounts of the period which we cite below.

4. The information in parentheses identifies the speaker.

5. The chronology of ABE copula research is not reflected by its publication dates; Labov (1972) was reported in 1967 and first published in 1969; Holm (1984) was first reported in 1976; Rickford's and Baugh's papers were written quite some time before they were published.

6. Our analysis was replicated by Sali Tagliamonte on four interviews taped in 1982 with informants not in the 1981 sample. The results, again based on more than 500 tokens, were virtually identical to those we report here, particularly regarding adjectival influence on contraction and deletion.
7. Neither speaker 4 nor speaker 8 produced many tokens of the copula during joint recording sessions with their respective spouses (speakers 3 and 7), so henceforth their data will not be considered separately.

8. Most authors cited provide quantitative documentation of the behavior of is only following the pronouns he/she and full NPs. Even our low estimate of 0.15 for proportion of deleted copula with NP subjects must be considered somewhat inflated since it is based on data containing a few plural NPs: Are is much more readily deleted.

9. In a context with subject X, preceding segment Y, and following category Z, the probability of contraction is estimated to be \( p = \frac{m}{1 - p} - \frac{m \times x}{1 - x \times y} = \frac{1 - y \times z}{1 - z} \), where \( m \) is the corrected mean, and \( x, y \), and \( z \) are the estimated contributions shown in the table for features \( X, Y \), and \( Z \) respectively.

Because of the paucity of data on wh-clauses, this category would have shown categorical contraction if it were not combined with the NPs. The true NP figure would be even lower.

10. Wolfram (1974) found little or no effect of following consonant on are-deletion in white southern speech.

11. Singler has replicated our multiple regression analysis on the Liberian data with generally parallel results (where both data sets have sufficient tokens) with respect to constraint ordering for both contraction and deletion.

12. In comparison with child and adolescent varieties of ABE also studied quantitatively, Samaná English of course appears quite conservative. Thus for example, the Samaná proportion of deleted is after NP and PRO is .15 and .20 respectively. This contrasts with Labov's (1972) New York City Cobras (.31 and .72), Pfaff's (1971) Los Angeles children (.42 and .60), and Kovac's (1980) Washington, DC seven-year-olds (.27 and .48). However, such data is not entirely pertinent to our argument since the differences may be at least partly due to age-grading (Wolfram 1987) and may also be affected by ongoing linguistic change (Bailey and Maynor 1985).

REFERENCES


______. 1986b. "Liberian Settler English." Workshop on Creoles in Time, Space and Society, Linguistic Society of America Summer Institute, City University of New York.


