Don’t know much about history: Beyond language in antebellum African American correspondence

Gerard Van Herk, University of Ottawa

1. Introduction

The focus of sociolinguistics – investigating the grammatical systems underlying variation in the speech of everyday people – has led to the collection of large data sets of contemporary vernacular speech suited to quantitative analysis. But historical sociolinguists face a sort of preserver’s paradox: archivists tend not to keep the writings of ordinary people, even when ordinary people were lucky enough to learn to write; and ordinary people who did produce archivable material usually became literate in a variety so standard that their writing hid whatever non-standard features might have been part of their speech. Reliable historical information on vernacular African American English (AAE) is even scarcer, as it was illegal in every slave state at some point to teach slaves to read and write (Cornelius 1991).

In this paper, I sketch out some of the creative effort that my subdiscipline has invested in discovering sources of earlier AAE. I then describe the creation of the Ottawa Repository of Earlier African American Correspondence (OREAAC), designed to complement these earlier efforts while addressing their shortcomings. I illustrate our use of the OREAAC to uncover previously-inaccessible information about regional variation in 19th-century AAE, and conclude with some speculation about potential further applications of (and improvements to) the corpus. At each stage of this process, I invite readers to consider the utility of a more interdisciplinary approach to our work, in both directions:

• What techniques and data sources could historians contribute to the production and evaluation of information on early AAE?

• In what way could existing sociolinguistic corpora, especially the OREAAC and any future iterations thereof, serve as data for historical work beyond the ability or interests of linguists?

2. Sources of information on early African American English

Despite the great volume of research on AAE, we are still unsure where the contemporary language came from and how it may have changed over the centuries. We know even less about how AAE may have varied from one location to the next. A more complete picture of earlier regional variation might help us better understand the development of both the language and the variation that persists to this day.

Much recent work has speculated on the likely linguistic outcomes of sociohistoric factors such as the period of settlement, African-European population ratios, and type of agriculture (e.g. Mufwene 1996, 2000; Rickford, 1997; Winford, 1997, 1998; Wolfram & Thomas, 2002). In particular, distinctions have been drawn between the deep south states of Louisiana, Mississippi, Alabama, Georgia and South Carolina, with their larger plantations and larger African-origin populations, and the remaining southern states -- the middle south. However, while historical evidence can serve to inform linguistic reconstruction, it can never give us definitive answers about the linguistic consequences of a particular situation (Van Herk, 2002:14; Walker, 2000:23-4). To address those issues, we must supplement such information with linguistic evidence (if we can find it).
The past half century has seen a great deal of creative effort invested in discovering sources of earlier AAE. Usually, these sources are evaluated based on criteria relevant to the study of the variety as a whole. Are the data primary, actually produced by African Americans, or secondary, descriptions or transcriptions by others? Are the data historical, actually produced during the time period of interest to us, or semi-contemporary, recently produced but in contexts where we would expect persistence of earlier linguistic systems? If sources are historical, how far back do they take us?

However, even the best of these sources are not ideally suited to the investigation of regional variation in early AAE, which requires large amounts of primary historical data and detailed information on the origin of informants. Recorded interviews with ex-slaves (Bailey, Maynor, & Cukor-Avila 1991) and early written materials (Montgomery 1999) are too sparse. Greater volumes of appropriate data can be gathered from sociolinguistic interviews with descendants of 18th- and 19th-century African Americans who settled in linguistically isolated communities in Nova Scotia, the Dominican Republic, and Liberia (Poplack & Tagliamonte 2001, Singler 1989). Unfortunately, the original settler groups cannot always be attributed a precise place of origin (and the language may involve post-settlement dialect leveling). Printed ex-slave interviews (Rawick 1972, Schneider 1989), while permitting regional analysis, cannot by the nature of their transcription process capture fine details of linguistic significance.

3. Building and evaluating the OREAAC

These concerns led to a project designed to construct a corpus of Early AAE which is both primary and historic -- a compilation of 427 letters written between 1834 and 1866 by African American immigrants to Liberia, both manumitted slaves and freeborn Blacks from all areas of the United States. Our team at the University of Ottawa chose these letters from the archives of the American Colonization Society (ACS), founded in 1816-1817 to encourage the removal from the U.S. south of the free African American community. The ACS managed to send over 15,000 African Americans to Liberia. Their letters, from Liberia and the United States, form part of the collection of 191,000 surviving documents housed at the Library of Congress in Washington. Twenty-nine of these letters were published by historian Bell Wiley in 1980, but the great majority were never transcribed or analyzed prior to our research. The development of the corpus, in particular the focus on finding the most speech-like letters by barely-literate non-elite settlers, is described in detail in Van Herk & Poplack (2003). The density of AAE pronunciation and grammatical features in the letters is illustrated in (1).

(1) thar wors my Sister Nan that die Sriker the agin toke all of her money and left the Child thar an orhans af the is any law for that I watter nor et my morer is alive fur sister Six boys all alive tok the child money to and wontter sel the things Bot we wontt consent So all what She fish  with her w have now

‘There was my sister Nan, that died. Striker the agent took all of her money and left the child there on our hands. If there is any law for that, I want to know it. My mother is alive, her sister’s six boys [are] all alive. [He] took the child’s money too, and wanted to sell the things, but we wouldn’t consent. So all that she fetched with her, we have now.’
Despite the potential value of correspondence, only a few scholars have looked at this type of data. Material is difficult to find; many researchers question the value of African American letters as representative of AAE speech, and of literate African Americans as representative of African American society. The common assumption, that literate African Americans represented a small and radically distinct subset of African American society, is based on two widely-held beliefs: that literacy was solely the domain of house servants taught by their masters, and that few African Americans were literate. Both those beliefs seem to be challenged by recent historical research, which situates the transmission of literacy squarely within the African American community itself. Anderson (1988) characterizes the “typical experience” in the words of former slave Louisa Gause, as in (2).

(2) “No child, white people never teach colored people nothin, but to be good to dey massa en mittie, what learning dey would get in dem days dey been get it at night; taught themselves.” (Anderson 1988:17)

And actual literacy rates for antebellum African Americans simply aren’t known. The widely-quoted estimate of 5 to 10 percent overall is based on informants in the Ex-Slave Narratives (Rawick 1972) who mention being literate during slavery, and the 25% figure for Liberian immigrants is based on census data and ships’ logs (Shick 1971, Brown 1975). But in both cases, the number of illiterate informants has never actually been tabulated; the remaining respondents simply didn’t mention literacy.

<table>
<thead>
<tr>
<th>Ascribed degree of literacy:</th>
<th>( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>unknown</td>
<td>81</td>
</tr>
<tr>
<td>spells</td>
<td>2</td>
</tr>
<tr>
<td>reads</td>
<td>16</td>
</tr>
<tr>
<td>writes</td>
<td>2</td>
</tr>
<tr>
<td>reads and writes</td>
<td>13</td>
</tr>
<tr>
<td>good education</td>
<td>2</td>
</tr>
<tr>
<td>Total correspondents identified in documents</td>
<td>116</td>
</tr>
</tbody>
</table>

Sources: Liberian roll and 1843 census (Shick 1971, Brown 1975)

Table 1 demonstrates that most of the OREAAC correspondents for whom we have historical documentation fall into this “unknown” category – in other words, these letters are written by the same people generally assumed to be illiterate! The second largest group of correspondents are those who claim to read, but not write. Perhaps (basic) literacy may have been more prevalent, across a wider spectrum of antebellum African American society, than previously suggested.
Table 2. Reported demographic characteristics (in percentages) of OREAAC correspondents, Liberian settlers, and the African American population in the USA.

<table>
<thead>
<tr>
<th></th>
<th>OREAAC</th>
<th>Liberia, c. 1843</th>
<th>USA, c. 1850</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported illiterate</td>
<td>76.4</td>
<td>75</td>
<td>90-95 (slave)</td>
</tr>
<tr>
<td>Slave</td>
<td>60.9</td>
<td>59.9</td>
<td>86.6</td>
</tr>
<tr>
<td>Southern states</td>
<td>90.1</td>
<td>92.2</td>
<td>94.9</td>
</tr>
<tr>
<td>Agricultural</td>
<td>60.6*</td>
<td>37-74.7*</td>
<td>88.7</td>
</tr>
</tbody>
</table>


Liberian census and ship’s records let us address another question: Who were the OREAAC correspondents? How representative were they in terms of slave status, geographic origin, and occupation? As table 2 shows, most correspondents for whom we have information had been slaves in America, came from slave states, and appear to have been agricultural laborers. On all these axes, the majority of the OREAAC correspondents match the demographic characteristics of the majority of African Americans of the time; of greater importance, this information allows us to investigate the linguistic correlates of demographic characteristics.

4. What the OREAAC can tell us about regional linguistic diversity

Obviously a written corpus of non-standard language can only complement existing data sources, rather than replacing them. Despite this, the OREAAC has several strengths that make it ideally suited for investigation of regional variation: the volume of data permits quantitative analysis of multiple potential conditioning factors; the speech-like writing permits phonological analysis; and we can assign a state of origin to most OREAAC letter writers. While it is unlikely that 19th-century state and linguistic boundaries match perfectly, this information does permit broad general characterizations. Informants from Louisiana, Mississippi, Alabama, Georgia, and South Carolina were coded "deep south"; those from Virginia, Tennessee, Kentucky, North Carolina, and Washington, DC were coded "middle south." In keeping with the requirements of comparative sociolinguistics (Tagliamonte 2002), I focus less on simple frequencies of use of particular linguistic forms in particular regions; instead I compare the underlying systems of each region by analyzing quantitative relationships between each form and linguistic factors like verb type, subject, relationship between events, and pronunciation. Here, I describe briefly several analyses of variable verb marking in the OREAAC. These features are well-suited to a regional variationist analysis: they are frequent enough to permit study; a number of authors have found that they can remain variable even in writing (Funkhouser 1973, Montgomery 1999); and enough literature on their linguistic behaviour exists to delineate -- and possibly differentiate -- underlying linguistic systems. These analyses involved the extraction of over 8000 sentences from the OREAAC. Each was coded for up to a dozen linguistic factors proposed in the literature, all of which could be analyzed statistically at the same time with the help of a variable rule programme written specifically for linguistic analysis (Rand & Sankoff 1990). Details are available in Van Herk (2004) and Van Herk & Walker (2005).

4.1 Past tense marking on regular verbs…
Verbs that form their past tense form by adding an –ed suffix (e.g., *jumped*) often surface bare (e.g. *jump*) in AAE, as well as in many English dialects and creoles. The OREAAC features many such bare verbs, as in (3).

(3) on monday 18th I went up to Bexley and joined Mr Clarke (155/4/59)

However, the linguistic factors associated with bare forms (in sociolinguistic jargon, the *favouring factors*) vary across dialects and creoles. In the OREAAC, the strongest favouring factor by far is the presence of one or more consonants before the potential –ed slot – in other words, many OREAAC correspondents write *jumped* as *jump*, but *died* almost always stays *died*. This matches findings for the spoken language of African American diaspora communities in Nova Scotia and the Dominican Republic (Poplack & Tagliamonte 2001), as well as contemporary AAE everywhere (e.g., Labov et al. 1968). The correlation is true for both middle south and deep south writers, but the effect is much stronger in the deep south.

4.2 … and on irregular verbs

Irregular verbs, those that form their past tense in other ways (e.g. *come/came*, *go/went*), also show strong regional conditioning. In both the middle and deep south, bare forms are associated with a small group of verbs that have surfaced bare in English dialects for centuries: *come, run, give, and become*. Again, this matches both diaspora and contemporary AAE findings (Rickford 1999), and is parallel in both the middle and deep south. This time, however, it is in the middle south that the effect is much, much stronger.

4.3 Present tense s-marking

Another mainstay of research on AAE is the variation between zero marking, as in (4a), and -s marking, as in (4b), in verbs referring to present events.

(4)  a. My wife send her love to you both (159/9/153)  
   b. we Labors under a great deal of disadvantages (158/8/5)

The factors associated with this variation in contemporary AAE are complex and often contradictory. In the diaspora communities, though, the correlations are much clearer. S-marking is more common when describing longstanding states or actions, as in (4b); bare forms are more common when describing single events, as in (4a). In addition, s-marking is strongly associated with a complex interaction of third-person plural subject type and how close the subject is to the verb (don’t ask). This particular combination of factors has been operative in northern dialects in Great Britain for hundreds of years, and so is known as the Northern Subject Rule (NSR).

In the OREAAC, both habitual effects and the NSR are in effect. Again, though, we see a regional distinction – habitual effects are stronger in the deep south, while the NSR is stronger in the middle south.

All these findings together start to illustrate a much more nuanced and complex scenario for earlier AAE than many earlier (and highly controversial) accounts would suggest. The linguistic system of middle south OREAAC correspondents maintains strong grammatical effects from English dialects – and, in fact, these are the areas with higher white-black population ratios, and smaller landholdings, social factors that would
encourage transmission of dialect features. The linguistic system of the deep south, on the other hand, shows stronger evidence of phonological (pronunciation) effects and a habituality finding, both of which are potentially universal tendencies, or derived from creoles, dialects, or African languages. This area – the “Black Belt” -- featured lower white-black population ratios and larger plantations. The language of both regions is affected by both types of grammars, but in different proportions consistent with their demographic characteristics. This scenario only reveals itself through the combination of primary historical data and demographic information on correspondents found in the OREAAC.

5. Moving beyond

Extended research in this area is hampered by limits on the material available, but a few avenues of further research (linguistic and otherwise) do suggest themselves. The imposition of the Civil War as a temporal boundary on early AAE research was originally intended to test hypotheses whereby postwar literacy dramatically changed AAE; as the OREAAC correspondents are by definition somewhat literate, that limitation becomes irrelevant, and it would be useful to know more about what happened to AAE between the Civil War and about 1900, dates for which letters are available in the ACS archives. The immediate postwar period is marked by major Liberian immigration from Georgia and South Carolina; as many residents of these states would have been Gullah speakers who learned to read from other African Americans in the Freedmen's Schools, letters from these informants might shed light on the relationship between earlier stages of Gullah and AAE. It would also be useful to re-investigate archives for additional letters from states where material is in short supply, to perhaps permit finer regional distinctions (especially between Gulf and Atlantic coastal areas of the deep south). In the existing OREAAC, some other features, especially purely phonological ones, remain to be analyzed. Even findings to this point, however, allow us to begin propose a more nuanced model of early AAE, one revealing multiple roots, varying in their contributions across region, time, and linguistic domain, consistent with many of the suggestions from historical and language contact research.

In a less strictly linguistic vein, the OREAAC’s focus on non-elite correspondents makes it potentially useful as a source of voices that have not been heard in the story of 19th-century American and West Africa. The OREAAC features firsthand accounts of battles with native groups (5), including grim details censored from published versions in the ACS’s house organ, the African Repository.

(5) Gotorah returned, back to the kitchen, which he siesed and shook with one hand, and brandished a dreadfull knife with the other abot 6 inches broad, and about a hundred and 50 men come up to the fence, to whom he said let us go in, I took deliberate aim at him (he was half bint shaking) and brought to the ground cut off his knee shot >him] in the lungs, {strikeout:and} {strikeout: cut of his privets} (154/2/37)

Although the majority of letters (approximately 85%) are by male heads of households, the letters also feature “back-channel” communications between women in Liberia and America, as in (6).

(6) Mrs. S. McLan,
Dear Sister, I hope you are well & children. I had some feavor after I got home, But are well. I dream of seeing you often, & would like to see you, But I could not concnet to come to america. My love to you & all the children. how large is John I love him. I wish I could see him, when he grows up, he must send me a present. my respects to Revd. Mr. Mclane yours Martha A Harris (154/3/137)

The letters also often reveal a profound dissatisfaction with the administration of the ACS, both in Liberia (as in (2) above), and in America (7).

(7) when i Say to you My Dear Bro that i would Sue for our Rights i mean i would Sue the old Man heirs & Mr Mlain Because i think Mr Mlain know Sumthing about the Money that was left for us. (159/9/110)

Correspondents often show a streak of homesickness (8a), often revealed in requests to purchase American foodstuffs (8b).

(8) a. I am well pleased with the country, save the absence of the Flesh-pots which I miss very much. (160/13/154)
   b. always send as much beef and pork with all your expeditions for what is not consumed by the emigrants will always find reddy sale here (154/3/164)

On the other hand, correspondents also show enthusiasm for Liberia, both the land and the political entity (9).

(9) I sit down truly under my own vine and fig tree and I am happy & satisfied here (156/6/209)

In fact, we are attempting a semi-quantitative analysis of terms of ethnic and national self-reference in the OREAAC letters (e.g. coloured, black, African, Sons of Ham), which seems to be involved in a complex interplay of political and identity issues related to establishing the inherent right of African Americans to self-government (10).

(10) This is not much of a country or any great shakes of a people, but it is my country and my people, and it suits me very well. (160/13/154)

The major health benefit of corpus building, of course, is the development of 20/20 hindsight, and this is certainly true of the OREAAC. Our late discovery of demographic information sharply restricted our ability to establish full regional or social representativity (neither of which were, admittedly, our original intent in building the corpus). As well, our focus on the linguistic utility of the material led us to exclude letters that might prove useful if the corpus were expanded to serve cross-disciplinary purposes. And we weren’t assiduous enough in following up the suggestions from colleagues (notably Michael Montgomery of the University of South Carolina) that we involve historians in the process from the beginning, especially to develop a glossary of geographical and trade terms that would have expedited the transcription process. We hope that future iterations of the corpus, or sub-corpora, will benefit from this hard-won
knowledge, as well as the contributions of advisors and friends from across the disciplines.
References


*This research was made possible by a post-doctoral fellowship from the Social Sciences and Humanities Research Council of Canada, and a supplemental bursary from the University of Ottawa, for which I am most grateful.*
It is possible that some slaves may have been sold across regional boundaries. We are reassured that published and actual states of origin generally match by references in the letters themselves, as well as the fact that most of our informants are from states with long-standing slave populations. Work heavily based on regions that were destinations of early 19th-century population movements would have to proceed with greater caution. A similar caveat would hold for research on white southerners from those states (e.g. Preliminary investigation of a third feature, the use of the present perfect, revealed no regional distinctions whatsoever in use. This is not surprising, given that no dialectal or regional variation in the linguistic conditioning of this form is claimed in the literature.