

# 1

## *Introduction*

### 1.1 The Status of African American English

The synchronic and diachronic status of African American Vernacular English (AAVE) has now been scrutinized more than that of any other vernacular variety in the history of American English. In fact, a survey (Schneider, 1996:3) of published research on dialects of American English from the mid-1960s through the mid-1990s indicates that AAVE has had more than five times as many publications devoted to it as any other variety of English and more publications than all other varieties of American English combined. Given the level of attention, it is therefore somewhat surprising to find Singler's (1998) appraisal of historical research on AAVE highlight the paucity of data about earlier African American speech.

A vexing problem in determining the age of particular AAVE features has been the general absence of data about earlier stages. Moreover, the data that have been available have often been suspect because of the circumstances under which they were gathered, because of questions as to whether or not the speakers were actually speaking AAVE, and the like. (Singler, 1998a:227)

Despite periods of apparent consensus among sociolinguists, data on the origin and early development of AAVE are still quite limited, and debate over its genesis continues to be intense after almost half a century of inquiry.

The debate over the evolution of AAVE is hardly limited to its origin and early development. At the same time, there is continuing debate about its more recent development, particularly in relation to other vernacular varieties of English. Is AAVE changing in ways that make it more distinct from other vernacular varieties of English – the so-called *divergence hypothesis* (Labov, 1985; Fasold, 1987) – or is it aligning more closely with other varieties of

English? Arguments over the development of AAVE in the twentieth century are as contentious as the debate over its earlier history, even though the empirical data to address questions of language change in the twentieth century would seem to be more readily accessible than data related to earlier African American speech. Obviously, issues concerning the development of AAVE are not limited to the simple accumulation of data; they concern the reliability of the sources of data, the methods of analysis, and the interpretations of the results derived from different studies.

Our goal is to address some of the major issues in the historical and current development of AAVE by examining in detail a unique enclave dialect situation that we have uncovered in Hyde County, North Carolina, located along the coast of the Atlantic Ocean by the Pamlico Sound. Although the empirical study focuses on a single, long-standing biracial enclave situation in a remote geographical location, the implications for the more general origin and early development of AAVE are far-reaching. The data allow us to consider fundamental issues in the reconstruction of AAVE such as the effect of regional dialects on the earlier speech of African Americans, the role of persistent substratal influence from earlier language contact situations between Africans and Europeans, the nature of intracommunity language variation in earlier AAVE, and the trajectory of language change in earlier and contemporary African American speech. All of these issues are central to the resolution of the debate over the past and current development of AAVE.

## 1.2 A Unique Database

Data from two sources have fueled the re-examination of the earlier history of AAVE in the last couple of decades. First, there has been a significant expansion in uncovering written documentation representing earlier African American speech. Although previous accounts of earlier African American speech (Stewart, 1967, 1968; Dillard, 1972) often included citations from written records by African Americans and observations about African American speech, the detailed analysis of different types of written texts have challenged the textual reliability of some earlier written records, and correspondingly, the view on the early development of African American speech. One important type of data that came to light in the 1980s was a set of written records and audio recordings of ex-slaves. These include an extensive set of ex-slave narratives collected under the Works Project Administration (WPA) (Schneider, 1989) in the 1930s, letters written by semiliterate ex-slaves in the mid-1800s (Montgomery et al., 1993; Montgomery and Fuller, 1996), and other specialized collections of texts, such as the Hyatt texts – an extensive set of interviews conducted with black hoodoo doctors in the 1930s

(Hyatt, 1970–8; Viereck, 1988; Ewers, 1996). All of these records seemed to point to the conclusion that earlier AAVE was not nearly as distinct from postcolonial Anglo-American English varieties as earlier hypotheses had proposed, namely, those hypothesizing a protocreole origin for AAVE (Stewart, 1967, 1968; Dillard, 1972). Although emerging written documentation on earlier African American English has raised essential issues related to reconstructing African American speech (see chapter 2), we have little to add to this discussion, since we have uncovered no written records from earlier African Americans in Hyde County.

The second type of database that inspired the re-examination of the historical development of earlier African American speech came from the investigation of black expatriate varieties of English. For example, in the 1820s, a group of blacks migrated to the peninsula of Samaná in the Dominican Republic, living in relative isolation and maintaining an apparent relic variety of English up to the present day (Poplack and Sankoff, 1987; Poplack and Tagliamonte, 1989; Poplack, 1999). A significant population of African Americans also migrated from the United States to Canada in the early 1800s, and some have lived to this day in relative isolation in Nova Scotia. The examination of the English varieties spoken by blacks in these areas by Shana Poplack and Sali Tagliamonte (1991), and their team of researchers (Poplack, 1999; Poplack and Tagliamonte, 2001) has led to the conclusion that these insular varieties were quite similar to earlier European American varieties, again raising important challenges to the hypothesis that a protocreole language was implicated in the origin and early development of African American English. The validity of such evidence is premised on three assumptions: (1) that the transplant language variety of the expatriates was an authentic reflection of a vernacular variety typical of many African Americans at that time; (2) that such communities would be relatively conservative in their language change in their new settlement communities; and (3) that these communities would remain relatively unaffected by the influence of surrounding communities and immune to changes taking place in contemporary AAVE. There are important questions that have been raised with respect to each of these issues, not unlike the kinds of questions we must confront in our study of Hyde County.

The community we examine here falls squarely within the tradition of enclave dialect studies, though it is a different kind of situation in that it involves a long-standing, relatively isolated biracial community in the rural Southern United States. Although expatriate transplant communities such as those studied by Poplack, Tagliamonte, and their team of researchers at the University of Ottawa (e.g., Poplack, 1999) may seem, at first glance, to hold more potential for examining the state of earlier African American speech than the coastal North Carolina Hyde County community because of greater physical dislocation and obvious social detachment from cohort

communities of European Americans in the USA, we would argue that enclave situations in the USA may offer equally compelling insight. In the community we examine here, there is comparable geographic remoteness and social detachment though the physical dislocation may not be nearly as great as that involved in expatriate situations. There is, however, continuity within the community for a much longer time frame than that found for transplant situations, since the community is now almost three centuries old.

As we will see in our detailed description of Hyde County in chapter 4, it is one of the oldest European American and African American settlement communities in North Carolina. Europeans settled there at the turn of the 1700s, and shortly thereafter African Americans were brought to the area. After a period of growth in the early 1700s, the area became an isolated enclave located in terrain that was 85 percent marshland. This setting, which has maintained an African American population of between a quarter and a half of the overall population of Hyde County throughout its history, provides an ideal setting for examining several critical issues regarding the historical development of African American speech. For one, it offers a sociolinguistic context involving a long-term, relatively insular, biracial situation featuring a distinctive European American variety. The Outer Banks dialect described in a number of recent publications (e.g., Wolfram and Schilling-Estes, 1995, 1997; Wolfram, Hazen, and Schilling-Estes, 1999) is found in this mainland setting as well as in the Outer Banks. The location of Hyde County and the approximate extent of the Pamlico Sound dialect area are given in figure 1.1.

The historical continuity of the African American community in the Hyde County region – almost three centuries old now – also provides an important perspective on the possible genesis and early development of AAVE. For example, the family genealogies of many of the current European American residents date back to the earliest residents of the area.

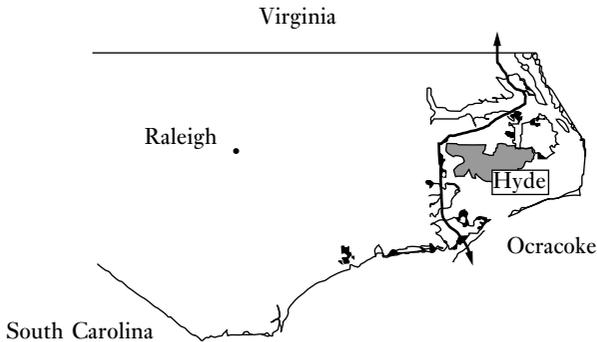


Figure 1.1 Location of Hyde County and the traditional Pamlico Sound dialect

Although comparable genealogies do not exist for African Americans, oral histories and emerging genealogies (see chapter 4) point to a parallel historical continuity. This observation is confirmed by the large number of shared surnames found among African Americans and European Americans in Hyde County – naming practices that reflect the tradition of the antebellum South when slaves were assigned their owners' surnames. The relatively stable census figures of Hyde County over two centuries of official census records also reflect the fact that there has been limited in-migration into Hyde County.

Finally, this situation offers insight into how African American speech is presently developing with respect to local European American vernacular varieties of English, as well as to varieties spoken by African Americans elsewhere. Through the application of the apparent time construct, we can see how the residents of Hyde County have been changing their speech during the twentieth century. By comparing the trajectories of language change for African Americans and European Americans both in relation to each other and in relation to external norms we can provide answers to the prominent debate about ethnolinguistic accommodation and diversity in the twentieth century.

### 1.3 The Hyde County Corpus

Since 1997, the staff of the North Carolina Language and Life Project has interviewed 144 lifetime residents of Hyde County, comprising 92 African Americans and 52 European Americans, as part of an ongoing sociolinguistic investigation. The age range of the speakers spans a century in apparent time, including those born as early as the 1890s to those born in the 1990s. Subjects for this study were chosen following the social network procedure of locating a friend of a friend (Milroy, 1987) and a family tree social network in which different members of extended families were selected for interviewing (Green, 1998). The use of the family tree procedure in selecting subjects was related to an effort to trace genealogies for some of the long-standing African American families in the community. The family tree sample offers the obvious advantage of comparing speech across different generations within the same family, while the social network sample offers the advantage of a broader representation of socially connected speakers within the community.

Some of the interviews followed the format of the traditional, conversationally based sociolinguistic interview (Labov, 1966; Wolfram and Fasold, 1974), while other interviews followed Green's (1998) genealogical inquiry interviewing procedure that focused on discussions of family lineage. Both

interviewing techniques led to extended conversations and significant amounts of casual speech data, although there may have been some consequences in terms of the conversational themes and the occurrence of particular structural features due to the differential topic focus in the interviews. For example, the two techniques precluded a systematic investigation of lexical variants, including such local terms as *pone bread* “cornbread made with molasses (treacle),” *juniper* “Atlantic white-cedar (*Chamaecyparis thyoides*),” and *shivering owl* “screech owl.”

Our analysis is based on data from the conversational portion of the interview, which typically lasted 60–90 minutes. A few of the speakers in our corpus were interviewed several times over the course of our fieldwork, so that we have three to four hours of conversation for these speakers. Both European American and African American interviewers of different ages and both sexes conducted interviews, although we do not consider the effects of interviewer variables such as ethnicity, sex, and age in the current analysis. Subjects were interviewed in their homes, at work, or at various meeting places where they spend their leisure time. Our goal was to interview subjects in a social setting in which they were most comfortable and relaxed. Generally, one or two members of our research team conducted the interviews, sometimes with one participant but in some cases with more than one subject. In some cases, a community member also accompanied the fieldworker, especially for interviews with members of the African American community. The majority of conversations are three-way conversations rather than dyads, following the model we established in the earlier study of Ocracoke (Wolfram, Hazen, and Schilling-Estes, 1999:6). We found that most participants felt more comfortable in a three-way conversation than they did in a one-on-one question and answer format.

Although we ended up with a convenience sample of Hyde County that followed social network and family network relations, attention was given to the representation of different age levels in our sample. We also sought to have adequate representation of men and women for both ethnic groups and different age groups, although we do not examine the role of gender in a systematic way as a part of this study (see Beckett, 2001). Therefore, our sample includes speakers who represent all ages for both ethnic groups. Such a generational distribution is critical for the goals of this study as they relate to language change and ethnolinguistic alignment in apparent time.

## 1.4 Data Analysis

The analytical procedures generally fall in line with the current methods used in quantitative sociolinguistics, with a couple of notable exceptions.

For morphosyntactic structures, discussed in chapter 5, a sample of diagnostic structures was first selected for detailed examination. The representative variables selected here include two morphosyntactic patterns typically associated with the Pamlico Sound regional variety, past tense leveling to *meren't* (e.g., *It weren't me*) and 3rd plural *-s* marking (e.g., *The dogs barks*), and two features commonly associated with the core morphosyntactic structures of AAVE, copula/auxiliary absence (e.g., *She nice*) and 3rd sg. *-s* absence (e.g., *The dog bark*). In comparing patterns of variation and change for a sample of AAVE-exclusive structures and distinctive Pamlico Sound morphosyntactic structures across different generations of African American and European American speakers, we hope to ascertain the ways in which the vernaculars were aligned at an earlier period in their history and how this alignment is changing. Accommodation to distinct Pamlico Sound morphosyntactic features by African Americans would show that they were sensitive to regionalized dialect norms. On the other hand, the persistence of dialect features associated exclusively with the African American speech community would give us insight into the maintenance of long-term ethnolinguistic distinctiveness.

The morphosyntactic analysis is based on a subsample of 49 speakers: 35 African American speakers divided into four age groups of speakers (12 young, aged 14–23; 6 middle-aged, aged 32–43; 6 senior, aged 55–70; and 11 elderly, aged 77–102) and 14 European American speakers divided into an elderly (6, aged 77–92) and a young (8, aged 15–27) group. The age delimitation for the four different groups of African American speakers follows the clustering of generational age groups in our sample of speakers (elderly = 77–91; senior = 55–70; middle = 32–43; young = 14–23) rather than the arbitrary, chronologically based increments sometimes found in studies of apparent time (e.g., 21–40, 41–60, etc.). To some extent, divisions between age groups also correspond to some significant historical events. For example, the two oldest groups of speakers represent speakers born during or before World War II, an event sometimes correlated with significant language shifts in the South (Bailey et al., 1991, 1993). Speakers in the two oldest age groups also attended segregated schools, whereas the two younger groups of speakers were educated during or after school integration took place (see chapter 4) in the Hyde County schools.

Because European American speakers serve primarily as a baseline group for comparison with the African American speakers in this study, samples of European Americans tend to be more limited in this study. As noted above, only 14 European Americans divided into two age groups – elderly speakers and younger speakers – were used in the morphosyntactic comparison and analysis of diagnostic consonants. Fewer European Americans than African Americans were used for the vowel analysis as well. It should also be noted that the European Americans selected for the morphosyntactic

analysis were limited to those classified as primary speakers of a vernacular version of Pamlico Sound English, based on independent linguistic criteria and background sociodemographic data. Judgments of vernacularity were made on the basis of sociodemographic background information as well as independent linguistic variables. All speakers in the sample who used negative concord and/or vernacular irregular verb patterns such as participial past tense forms (*I seen it*) or bare root past irregular forms (*Yesterday they come there*) were considered to be vernacular for the selection of this sample.

For the morphosyntactic analysis and the impressionistically based analysis of phonological variables (namely consonants), each instance of the diagnostic structure was extracted for the subsample of speakers, and the incidence of variants for each variable was tabulated for each speaker in the sample. However, in many cases, our analysis is based on summary descriptive data figures given for aggregate groups classified on the basis of ethnicity and age. Following the tradition of current quantitative sociolinguistics, we subjected these figures to VARBRUL analysis (Cedergren and Sankoff, 1974; Young and Bayley, 1996). VARBRUL is a probabilistic-based, multivariate regression procedure that shows the relative contributions of different factors to the overall variability of fluctuating forms. Factor groups may consist of independent linguistic constraints, such as following phonological environment, or external social ones, such as age group or social affiliation. The weighting values range from 0 to 1; a value of greater than .5 favors the occurrence of the variant, while a value of less than .5 disfavors its occurrence in a binomial analysis.

One of the limitations of VARBRUL analysis is its reliance on aggregate data that typically ignores the role of individual variation. But the relation of the individual to the group in sociolinguistics is hardly a settled issue. Thus, we also consider profiles of individual speakers to examine the role of the individual and group in sociolinguistic variation. In fact, one full chapter (chapter 9) is dedicated to the examination of the role of the individual variation in Earlier African American English, and much of our analysis of vowels focuses on individuals rather than groups. In this way, we hope to balance the analysis by considering both group and individual dynamics.

Chapters 6 and 7 treat variation in vowels and consonants, respectively, with chapter 8 devoted to intonation. The vowel analysis in chapter 6 comprises three parts. The first part consists of a general discussion of the vowel variants that make the dialect of Hyde County and the rest of the Pamlico Sound area distinctive. This discussion centers on a set of representative vowel formant plots. A detailed discussion of the acoustic measurement techniques, selection of tokens, and so forth used to produce these plots is provided. Next, we provide a historical survey of Hyde County vowels that compares our data with records from the *Linguistic Atlas of the Middle and South Atlantic States* (LAMSAS). The historical survey attempts to project

backward to reconstruct what Hyde County vowel configurations, both African American and European American, were like around 1800. In our estimation, the evidence from LAMSAS and from historical information about eastern North Carolina – especially about Hyde County, as described in chapter 4 – is adequate to make these conjectures. However, we do not feel that we can project back before 1800 with much confidence. Finally, we embark on a detailed acoustic analysis of twentieth-century developments in Hyde County vowel configurations. This analysis is based on a sample of 49 Hyde Countians, comprising 27 African Americans and 22 European Americans, who represent the entire range of ages interviewed for the study. It focuses on three vowels for which adequate numbers of tokens were available in the recordings: /ai/, as in *tide*; /au/, as in *out*; and /o/, as in *coat*. Using a linear regression analysis of the normalized mean formant values for these speakers, we show how African American and European American vowel configurations in Hyde County are changing with respect to each other and note the implications of these developments. Taken together, chapter 6 provides an overview of Hyde County vowels over a 200-year period, which allows us to reach broader conclusions about the long-range relationship of AAVE with European American vernaculars than most other studies have been able to do.

Our analysis of consonants (chapter 7) involves both a general qualitative profile of the phonological traits of African American and European American Pamlico Sound English and the detailed quantitative examination of two consonant variables, namely syllable-coda consonant cluster reduction and postvocalic *r*-lessness. The former variable is probably the most analyzed phonological variable in current quantitative sociolinguistics, but this replication offers insight into its historical role and current development within an insular biracial enclave situation. The other consonant variable we examine, postvocalic *r*-lessness, is also well represented in the sociolinguistic literature, including regional and ethnic varieties of English, but the regional pattern of postvocalic *r* in Hyde County is quite distinctive, thus making it particularly diagnostic in terms of the examination of historical and current ethnolinguistic alignment.

In contrast to the heavily studied variables of chapter 7, chapter 8 examines one of the most sparsely studied variables, intonation. We concentrate on one intonational variable, the frequency of high pitch accents. The ethnic alignment appears to be as strong for intonation in Hyde County as for other variables. However, the paucity of research on African American intonation prevents us from drawing any conclusions about the past state of AAVE intonation. Our inquiry in chapter 8 opens up at least as many questions as it resolves.

While our analysis hardly covers the entire spectrum of diagnostic structures for African American English and Pamlico Sound English, it includes

a fairly representative set of diagnostic phonological and grammatical features. In this respect, it is somewhat different from some of the current analyses of insular African American situations that tend to focus exclusively on a single feature or a restricted set of AAVE structures. Our investigation shows that an authentic picture of earlier African American speech can emerge only if a wide array of dialect structures is considered, including overall profiles of vowel systems, consonants, and prosody, along with diagnostic morphosyntactic variables. While selective, single-structure studies may reveal significant insight into a particular linguistic process and/or a dimension of an ethnolinguistic boundary, such studies may obscure or even distort our understanding of the overall relationship of African American speech to other varieties. We also complement some of our objective findings with results obtained from a speaker identification task. It is essential to keep both the objective and subjective dimensions in mind as we attempt to capture a clear picture of change and variation in earlier and contemporary African American English.

## 1.5 Beyond Hyde County

On the one hand, it may seem presumptuous to generalize about the earlier and the contemporary state of African American speech on the basis of a single community such as Hyde County. Certainly, we are mindful of the many cautions that must be exhibited in extrapolating from this situation to the general development of AAVE, regardless of how insightful this situation might be. On the other hand, there appear to be some unavoidable implications for the more general study of African American speech that derive from the examination of this unique sociolinguistic situation. These data allow us to address several critical issues related to the past and present evolution of African American speech in the USA.

One of the crucial issues addressed here is the accommodation of earlier African Americans to regional speechways. Few empirical situations can inform us as richly about the potential for such accommodation than the study of an African American community situated within a highly distinctive regional variety. At the same time, the persistence of long-term ethnolinguistic distinctiveness is a critical issue for reconstructing the history of African American speech. Recent studies of African American speech (Labov, 1998; Poplack, 1999) have suggested that the distinctiveness of AAVE is primarily a twentieth-century phenomenon, but our evidence suggests that this view may be somewhat exaggerated. In fact, our study suggests that, in many instances, earlier African American speech was more likely a mixture of regional traits shared with their European American cohort speech and

some persistent substrate features reflective of the earlier contact history between Africans and Europeans. Finally, our study addresses the issue of contemporary change, and the so-called divergence issue. Much has been made of the independent changes taking place within AAVE that are distancing it from comparable European American English vernacular varieties (Dayton, 1996; Labov, 1998). While there is selective support for some of these assertions, the most compelling evidence suggests that the crux of contemporary ethnolinguistic divergence may, in fact, lie in the development of a supraregional AAVE norm that entails the abandonment of, or resistance to, local regional norms. In this context, the results of our study are hardly limited to the small Hyde County community that is the focus of this study. Our findings should impact the core understanding of the earlier and present-day development of African American speech throughout North America in general.