# 28 Comparative Sociolinguistics

SALI TAGLIAMONTE

Comparison has always been at the root of sociolinguistics. The study of language behavior from a comparative perspective – comparative sociolinguistics – concerns the connection (relationship) of linguistic variation in one body of materials to another. This requires a methodology that, first, enables the many different influences on linguistic features to be disentangled through systematic examination of their behavior, and, second, that situates and explains the linguistic features through comparison with like features in related varieties. This methodology builds directly from two strands of linguistics – historical linguistics and quantitative sociolinguistics.

The comparative method in historical linguistics is based on comparative reconstruction, which has as its basis shared correspondences of linguistics features (e.g. Hoenigswald 1960, Meillet 1967). The application of these methods in sociolinguistics began with Weinreich et al.'s (1968) introduction of the notion of "structured heterogeneity" in the speech community which was later developed further by Labov (1982). This work laid the foundations of the quantitative variationist approach (Labov 1966, 1970, 1972, Labov et al. 1968) which elaborated a method of analysis founded on assumptions of accountability, testing hypotheses systematically against data, and building generalizations on well-formed comparative studies. Further, constraints on linguistic features were brought into the picture and were held to be a reflection of diachronic patterns even after centuries of geographic separation (Labov 1980: xvii).

A comparative approach had been, at least implicitly, adopted for tracking historical connections between related varieties since the turn of the century (Kurath 1928, 1964). Subsequently, it has been implicitly or explicitly adopted by numerous scholars in a wide range of applications: for making trans-Atlantic connections (Montgomery 1989a, 1989b, 1997); for contrasting data sets in real- and apparent-time (Bailey and Maynor 1985, Bailey et al. 1989, Cukor-Avila 1997, 1999); for tracing the roots of extraterritorial varieties of English (Clarke 1997a, 1997b, 1997c, Hickey forthcoming); and for isolating systems in language contact (Poplack and Meechan 1995, 1998a). More recently, a comparative

element has become increasingly prevalent in dialect studies (Wolfram 1999, 2000, Wolfram and Schilling-Estes to appear, Wolfram and Sellers forthcoming, Wolfram et al. 1997).

Comparative sociolinguistic research developed initially from issues surrounding the origins and development of African-American Vernacular English (AAVE) (Holm 1975, Rickford 1986, 1991, Rickford and Blake 1990, Singler 1991). This long-term debate provides a conundrum for the comparative sociolinguistic endeavor as researchers from all areas of the field attempt to reconstruct the likely characteristics of the ancestor of AAVE.

In this chapter my goal is to demonstrate, using case studies, comparative sociolinguistic methodology as it has developed in two research programs. The first involves tracking the origins and development of African-American Vernacular English (Poplack and Tagliamonte 1999, Poplack 2000, Poplack and Tagliamonte 2001). The second involves tracking the origins of nonstandard linguistic features of North American dialects in comparable British dialects (Godfrey and Tagliamonte 1999, Jones and Tagliamonte 2000, Tagliamonte and Smith 2000, Tagliamonte 1999).

I begin by describing the methods of the comparative sociolinguistic approach to language variation and change.

## 1 Variationist Sociolinguistics

Quantitative variationist methodology falls within the framework of empirical linguistics known as variation theory and employs multivariate analysis to model the linguistic phenomena under investigation, a type of analysis which forms part of the "descriptive-interpretative" strand of modern linguistic research (Sankoff 1988: 142–3). Studies employing this methodology are based on the premise that the features of a given speech community, whether morphosyntactic, phonological, lexical or discursive, may vary in a systematic way, and that this behavior can be quantitatively modeled (Young and Bayley 1996: 254). The approach rests on the assumption that whenever a choice exists among two (or more) alternatives in the course of linguistic performance, and where that choice may have been influenced by any number of factors, then it is appropriate to invoke statistical techniques (Sankoff 1988: 2).

The relationship between linguistic variables and external factors such as class and gender has often been criticized because social identities are not categorical or fixed notions, but are locally situated and constructed (Schiffrin 1996: 199). However, when the goal of research is to gauge and model the individual and combinatory effects of multidimensional internal linguistic factors alongside broadly defined external factors, a quantitative approach is particularly useful. The advantage of this type of analysis lies in its ability to model subtle grammatical tendencies and regularities in the data and assess their relative strength and significance when all possible factors operating on

them are treated simultaneously. The combination of factors exerting an influence on a given linguistic feature will often be extremely complex. The task for the analyst is to identify those factors which are the most meaningful for analysis and interpretation.

One of the foundations of variationist analysis is its attempt to discover not individual occurrences or overall rates of occurrence, but *patterns* of variability in the body (or bodies) of material under investigation.[1]

Practically speaking, we take the following steps.

1 Select an appropriate linguistic feature, ideally a diagnostic area of grammar:

   (i) circumscribe the variable context;
   (ii) code the data into factors which test hypotheses, claims and observations in the literature.

2 Examine the patterns of use of the linguistic feature according to the principle of accountability using the lines of evidence provided by statistical modeling techniques of multivariate analysis:

   (i) Which factors are statistically significant?
   (ii) What is the relative contribution of the linguistic features selected, i.e. which factor group is most significant (largest range) or least? (smallest range)
   (iii) What is the order (from more to less) of factors within a linguistic feature? (constraint hierarchy)
   (iv) Does this order reflect the direction predicted by one or the other of the hypotheses being tested?

Since linguistic change proceeds as "an ordered set of shifts in the frequency of application of the rule in each environment" (Labov 1982: 75) we can expect that not only rates, but especially the conditioning of linguistic variability will be language specific. Thus, the environmental constraints (i.e. the "factors" in item 1(ii)) on variation are the fundamental units of linguistic change (Labov 1982: 75), while the constraint ranking (i.e. 2(ii)) of factors provides a critical diagnostic for comparison. In this way, similarities and differences in the significance, strength, and ordering of constraints provide a microscopic view of the grammar from which we can infer the structure (and possible interaction) of different grammars. Thus, it is through the evidence provided by the various statistical techniques outlined in B above that we can "trace the path of linguistic development through a multidimensional space." These measures enable us to infer whether the data sets under comparison share an underlying grammar, and to what extent. For example, if the constraint ranking of one (or more) factor groups is shared by a set of varieties, we infer that they have inherited it from a common source. If the constraint ranking of factors is parallel, but operates at varying strengths or patterns in different varieties this

can be explained by the stage of development of the system of grammar under investigation as represented by each data set.

How are these procedures used to make comparisons and reconstruct origins? According to Poplack and Tagliamonte (1991: 318) determining the precise historical origins of a linguistic feature requires not only the existence of an apparently similar or identical feature in a putative source dialect, but also the same distribution in the language, as determined by the hierarchy of constraints conditioning its appearance. Thus, in order to determine the status of a form, it is not its current *existence* in a variety which is decisive, nor even its rates of occurrence. This is because overall rates of presence of absence of the variants under investigation will likely vary according to features of the situation (Poplack and Tagliamonte 1991: 318). However, the *distribution*, i.e. precisely where it occurs in the language, as determined by the relative frequency of the feature across its different contexts of use, is taken to represent the underlying grammatical structure.

The approach I describe here involves consistent comparison of each of the lines of evidence above but with the added triangularization of two or more relevant bodies of material to compare and/or contrast. This is where the comparative method comes in.

## 2 The Comparative Method

In historical linguistics it is widely held that earlier stages in the history of a language can be observed through comparative analysis of cognate forms (sets of reflexes) in later, sister varieties (e.g. Hoenigswald 1960: 119, Meillet 1967). The comparative method is "the procedure whereby morphs of two or more sister languages are matched in order to reconstruct the ancestor language" (Hoenigswald 1960: 119). In comparative sociolinguistics, the means by which the sister varieties are compared is the set of correspondences provided by the results of the statistical techniques of multivariate analysis, what Labov referred to as "finely articulated structures" (1982: 75). In fact, the quantitative paradigm provides the kind of "precise information on the states of the language" called for by Meillet (1967: 138).

We approach this by comparing the patterning of variability in each possible source. If the conditioning effects on the variable linguistic features show patterns approximating those found in a putative source, we can conclude that they represent structures drawn from that source (e.g. Poplack and Tagliamonte 1999, 2001). On the other hand, where there are dissimilarities, we have grounds for concluding that the phenomena in question belong to different linguistic systems (e.g. Tagliamonte 1998a, Tagliamonte et al. 1997).

A key notion in a comparative sociolinguistic approach is the notion "conflict site". This is defined as a form or class of forms which differs functionally and/or structurally and/or quantitatively across the varieties in question

(Poplack and Meechan 1998b: 132). By quantitatively analyzing patterns of distribution at grammatical sites where varieties are held to be distinct, the precise nature of similarities and differences across data sets can be pinpointed. If the results match the observations made in the literature for the putative source varieties, then we appeal to this similarity to posit a link between the two. On the other hand, where there are dissimilarities, we must contextualize and evaluate the differences in the context of linguistic developments in all the varieties under investigation. If the variability in the data is part of ongoing linguistic change, then it must be analyzed in terms of where it came from, consistent with Jespersen (1924) that "to understand a linguistic system, we must know how it came to be".

Thus, by utilizing the lines of evidence made available by variationist statistical techniques, we then proceed according to the following procedure:

3 Compare and contrast conditioning factors across sets of data which can be related (at least putatively) across some external set of criteria according to:

   (i) statistical significance;
   (ii) relative strength;
   (iii) constraint hierarchy.

I now illustrate this comparative sociolinguistic approach in practice and show how it can elucidate the nature of variability through an examination of similarities and differences across varieties.

## 3 Target of Investigation – Sisters under the Skin?

In the analyses that follow I will examine a number of different linguistic variables (diagnostics) in up to six different data sets. These data sets represent varieties which can be differentiated on a number of broad extralinguistic characteristics, as summarized in table 28.1.[2]

The data, which consist of hundreds of hours of tape-recorded conversations, include discussions about local traditions, narratives of personal experience, group interactions, and local gossip. Moreover, the interviews in BCK, GUY, NPE, and DVN were conducted by community members. All of these materials are highly informal and as far as possible represent the typical discourse found in each community.

Aside from the samples from OTT and YRK, the speakers in each of the corpora have relatively homogeneous socioeconomic characteristics. They were born and raised in the community in question and in each case they represent the oldest living generation at the time of the fieldwork. They are usually employed in traditional or service industries. Level of education among the

**Table 28.1** Extralinguistic characteristics of varieties under investigation

| Locale | Abbreviation | Geographic location | Separation from mainstream | Ethnic affiliation |
|---|---|---|---|---|
| Buckie | BCK | Northeast Scotland | yes | British |
| Guysborough Enclave | GYE | Nova Scotia, Canada | yes | African |
| North Preston | NPR | Nova Scotia, Canada | yes | African |
| Ex-Slave recordings | ESR | Southern United States | yes | African |
| Samaná | SAM | Dominican Republic | yes | African |
| Devon | DVN | Southwest England | yes | British |
| Guysborough village | GYV | Nova Scotia, Canada | intermediate | British |
| Ottawa | OTT | Ontario, Canada | no | British |
| York | YRK | York, England | no | British |

informants ranges from none to 12 years, with most speakers falling into the lower range. The speakers are similar in that they are members of "dense" networks (Milroy 1980) in that their social circles were generally confined to the community in question.

The communities are also differentiated by the ethnic ancestry of their inhabitants. In four data sets the speakers are of African descent (SAM, ESR, NPR, GYE); in the other four, the speakers are of British ancestry (GYV, BCK, DVN, OTT). Although all the African speakers represented in these data sets live in relatively isolated circumstances, the speakers of British ancestry represent a range of different backgrounds ranging from highly isolated (BCK) to mainstream/standard (OTT). Thus, we are provided with an unprecedented opportunity to conduct a cross-variety comparison in which linguistic features may be viewed across these two extralinguistic dimensions relatively independently.

Moreover, these contrasting extralinguistic characteristics along with standard accounts of the effects of language contact (Pousada and Poplack 1982, Thomason and Kaufman 1988), would lead us to expect that the more separate from mainstream culture, the higher the degree of impermeability to influence from surrounding mainstream vernaculars (see Poplack and Tagliamonte 2001). In the case of the Nova Scotian communities, these considerations, in conjunction with a standard diffusionist hypothesis, would lead us to expect that NPR should retain more local vernacular features; while GYE may show similarities with neighboring GYV. In the case of the British varieties which are situated at two extremes of dialect regions in Britain, one in the northeast (BCK) and the other in the southwest (DVN) we would expect both locales to retain local vernacular features, but that those features might be highly differentiated.

Each community differs with respect to the relative degree of exposure of residents to mainstream culture and language. OTT and YRK are clearly situated in the mainstream, while GYV, a rural village in Nova Scotia, stands in an intermediary position (see for discussion Poplack and Tagliamonte 1999). The remaining areas are rural and, in addition to geographic separation, they have all have had relatively limited contact with mainstream culture and outsiders.[3] In each, the speakers live in a remote fringe area, where they are also separated from large urban populations on additional sociocultural grounds. In these cases, each data set represents a variety which has evolved in a context of relative isolation.

These data bear many, if not all, of the characteristics of "peripheral" areas, which in historical linguistics are widely-known to provide choice evidence about earlier stages of a language (e.g. Anttila 1989: 294, Hock 1986: 442). Indeed, the two varieties from Britain can be characterized as highly conservative. Each retains features recorded in historical English, which have since become obsolescent or moribund. For example, BCK retains [f] in *wh-* words, velar fricatives, *-en* participles and adverb placement between verb and complement, while DVN retains initial voiced fricatives, pronoun exchange and *thee* as 2nd person pronoun. Many of these can be traced back to at least the Early Modern English period, with some having arisen in Old English and early Middle English supporting an interpretation of their status as peripheral.

Preservation of features from earlier stages in the history of English in these and other comparable communities are reported extensively in the sociolinguistic literature (e.g. Hazen 1996, Poplack and Tagliamonte 1989, Poplack and Tagliamonte 1994, Schilling-Estes and Wolfram 1994, Tagliamonte 1997b, Tagliamonte and Poplack 1988, Wolfram and Sellers forthcoming). Thus, these peripheral varieties can provide an interesting test site for models of language change and a critical window on the past.

In the remainder of this chapter, I present a series of analyses which consistently compare different combinations of these data sets, depending on the linguistic feature under investigation.

## 4 The Importance of Proportional Analysis

One of the most widely-studied areas of grammar to receive comparative investigation is variation between marked and unmarked past temporal reference weak verbs, as illustrated by variable marking on the verb *look* in (1):

(1) (a) Bunch of us walked up the stairs and sat down and Caroline *looked* up. (NPR/039/735–6) (Poplack and Tagliamonte 2001)
(b) When I *lookØ* in like that, and I *lookØ* in that door, and I *lookØ* back in the corner, I seen them great big eye. (NPR/030/884–6) (Poplack and Tagliamonte 2001)

**Table 28.2** Simple count: number of bare verbs with past temporal reference

| | NPR | GYE | SAM | ESR |
|---|---|---|---|---|
| Number of bare verbs | 127 | 159 | 518 | 100 |

*Source*: Meechan et al. (1996)

**Table 28.3** Total verbs in past contexts

| | NPR | GYE | SAM | ESR |
|---|---|---|---|---|
| Number of bare verbs | 127 | 159 | 518 | 100 |
| Total verbs in past contexts | 362 | 534 | 1234 | 283 |
| Percent of bare verbs (%) | 35 | 30 | 42 | 35 |

*Source*: Meechan et al. (1996)

This is an area of the grammar which is widely-agreed to differentiate Standard English, English-based creoles, indigenized Englishes and contact vernaculars (e.g. Patrick 1999, Winford 1992, Wolfram and Hatfield 1984).

Let us examine this feature in four of the varieties whose origins can be traced to a common geographic area (the United States), but which have been separated for 150 years in widely-separated contexts in Canada (NPR and GYE), the Dominican Republic (SAM), and the Southern United States (ESR).

First, let us simply count the number of unmarked (bare stem) verbs with past temporal reference in each body of materials, as in table 28.2.

This count of the data makes it look like one variety, SAM, has considerably more zero forms than the other corpora ($n$ = 518). This is, in fact, precisely what would be expected of creoles since these varieties are widely-known to have considerably less verbal affixation than other languages (e.g. Bickerton 1975). SAM is located in the Caribbean where many creoles are spoken and further, Samaná English is spoken by people of African descent. Since most creoles are spoken by people of the same ethnic affiliation, then one might be led to hypothesize that the reason SAM has many more bare verbs is because it has a more creole-like grammar than the other varieties.

However, bare numbers do not take into account the *proportion* of these verbs that occur of all relevant verbal constructions in each data set. Table 28.3 displays a distributional analysis according to the "principle of accountability" (Labov 1972: 72), in which the number of bare verbs is reported as a proportion of the total number of relevant constructions (either inflected or not).

It is now clear that the apparent differences between varieties observable in table 28.2 comes from the fact that there is a disproportionate number of bare

verbs in the data. The SAM data simply had more past temporal reference contexts. The percentages show that the overall rate of bare verbs across varieties is quite similar. The exact opposite of the result indicated by table 28.2.

This illustrates a fundamental component of the comparative sociolinguistic approach. It is necessary to deal with proportions in order to compare rates consistently and accountably across data sets.

However, as we shall see, even such a calculation provides only a first step in demonstrating that the varieties in table 28.3 are patterning in the same way. In fact, this view of the data reveals very little about the mechanism, i.e. internal organization, of the variability. Thus, it provides insufficient evidence which would enable us to distinguish between contrasting grammatical systems.

## 5 Contrasting Constraints across Varieties

As it happens, this linguistic variable, i.e. variation in marked and unmarked verbs in past temporal reference contexts, is a good conflict site to compare varieties because the environmental constraints on its application can be expected to differentiate grammatical systems. In creole vernaculars widespread and frequent absence of inflection on past reference verbs is said to be the result of an underlying stative/nonstative distinction combined with a tense/aspect system that is relative, rather than absolute. When past time is marked, the marker is there as an *anterior* marker, to mark the relationship of states and events in the discourse to each other. Even a variety at an advanced stage of decreolization (approximating Standard English norms) may still reflect this underlying grammatical organization of past markers. In Standard English, however, all events prior to speech time are required to mark past tense. The device most frequently employed for this is the preterite, where weak verbs take a suffix, i.e. *–t,d*. However, quantitative analyses of this feature in English dialects have revealed regular phonological conditioning on these word final suffixes, such that *–t,d* may be deleted by surface level (consonant cluster) reduction processes (e.g. Guy 1980, 1991, Guy and Boyd 1990, Neu 1980, Santa Ana 1996).

Exactly which constraints are in operation in the varieties depicted in tables 28.3 and 28.4? Is it consonant cluster simplification or a distinct grammatical system involving anterior marking coupled with a stative/nonstative distinction? Whichever ones they are will have a major bearing on making a decision about the underlying system (or grammar) in a given data set. Such a decision cannot be based on frequency or proportion alone. How can we decide on the nature of variability across varieties?

Let us now employ the techniques of variable rule analysis and the lines of evidence offered by statistical significance, relative strength of factors and constraint ranking of factors to consider the variation between marked and

**Table 28.4** Five independent variable rule analyses of the contribution of selected factors selected as significant to the probability that weak verbs will surface as *stems*

| | SAM | ESR | NPR | GYE | GYV |
|---|---|---|---|---|---|
| Corrected mean: | 0.45 | 0.29 | 0.31 | 0.59 | 0.14 |
| Total *n*: | 1,236 | 281 | 360 | 503 | 282 |
| *Preceding phonological segment* | | | | | |
| Consonant cluster | 0.81 | 0.73 | 0.73 | 0.62 | 0.76 |
| Single consonant | 0.60 | 0.51 | 0.55 | 0.55 | 0.70 |
| Vowel | 0.26 | 0.32 | 0.35 | 0.35 | 0.11 |
| *Following phonological segment* | | | | | |
| Consonant | 0.58 | 0.65 | 0.68 | 0.72 | 0.81 |
| Vowel | 0.38 | 0.32 | 0.31 | 0.29 | 0.30 |
| Factors not selected: | | | | | |
| Stativity/anteriority | X | X | X | X | X |

*Source*: Adapted from Poplack and Tagliamonte (2001)

unmarked verbs in table 28.4. Further to the comparison in tables 28.2 and 28.3, a critical comparative component is added – data from GYV, the variety of English spoken by the descendants of British loyalists in Guysborough, Nova Scotia (GYV), the closest village to GYE. This provides a sample which can be expected to embody conservative English patterns, a critical control for the other four varieties which involve isolated conditions and speakers of African descent.

Table 28.4 (and all tables in ensuing sections) presents all the information necessary for interpreting the variable rule analyses that have been performed on each data set. The *corrected mean* at the top of the tables indicates the overall tendency of the dependent variable (in this case the verb stem) to surface in the data. The *total n* records the denominator of the total number of contexts treated in the analysis. Each of the factor groups that have been considered in the analysis are listed with the results for each factor. Point-form numbers are *factor weights*. These indicate the probability of the dependent variable to occur in that context. The closer these numbers are to 1, the more highly favoring the effect is; the closer they are to zero the more disfavoring the effect is. The *range* indicates the relative strength of the factor. The higher this number is, the greater the contribution of that factor to the probability of the form.

We can now see that all five varieties behave near identically. Despite the varying corrected mean values which range from .14 in GYV, to .59 in GYE, the same phonological factors are chosen as significant and condition the variability

in the same way. In each case the ranking of more to less is parallel. *Preceding phonological segment* consonant clusters are the most highly favorable environment for stem verbs, then single consonants, while vowels disfavor. *Following phonological segment* consonants favor stem verbs but vowels disfavor. On the other hand the major factor relevant to creoles (stativity/anteriority) was not even selected as significant.

These results provide substantial evidence to conclude that the zero-forms on weak past temporal reference verbs in all these communities are the result of surface level phonological reduction processes, rather than the underlying functional distinctions of stativity and anteriority. Is this enough evidence to give a decisive answer to the question of the underlying system, i.e. what grammar underlies these varieties and what their source may have been?

The problem is that phonological constraints such as these which involve consonant cluster simplification of *–t/d* may be the result of universal phonotactic principles of grammar. If so, they can tell us little about the origin of these varieties. This highlights the fact that all linguistic features do not provide the same calibre of evidence for cross-variety comparison. Many so-called conflict sites are actually not *conflict* sites at all, since the same surface forms may appear across varieties that have no filial relationship. In other words, the constraints operating on the variation may be irrelevant to the issues of origins and system identification.

Let us now consider a linguistic feature whose variable *forms* appear globally, but for which examination of the historical record reveals distinctive patterning which can be traced to different source dialects in Britain. This obviates the possibility that the patterns in the data would have arisen independently.

## 6 Using Constraint Hierarchies to Disentangle Source Dialects

"Vernacular features" which appear robustly in dialects of English all over the world (Chambers 1995: 242–9), such as *was/were* variation in example (2), are particularly useful for the comparative sociolinguistic endeavor. The inter-variety parallelism in form permits consistent analysis of variable constraints which operate on their distribution and conditioning.

(2) (a) We *were* all thegither . . . I think we *was* all thegither. (BCK/h:72.44) (Smith and Tagliamonte 1998)
(b) He *was* lost all night once, and when he come back he *were* covered in dung. (DVN/001/56,43) (Jones and Tagliamonte 2000)
(c) There *was* a lot of us that *were* sort of seventeen. (004/180,27) (Tagliamonte 1998b: 155)
(d) It *weren't* us with the funny accent; it *was* them. (Schilling-Estes and Wolfram 1994: 298)

One explanation for *was/were* variation is that it is the result of "regularization" processes in language (Fries 1940). This is based on the idea that the verb *to be* is gradually becoming more like the other more regular verbs in English in having the same form, i.e. *was,* throughout the verbal paradigm, rather than the more complex distinction between 1st and 3rd person singular *was,* and 2nd person singular and 1st, 2nd, and 3rd person plural *were*.

However, the question that arises is where did the synchronic variation illustrated in (2) come from in the first place? One hypothesis is the standard diffusionist explanation, in which *was/were* was carried to these locations by people speaking varieties which contained the same features (Weinreich et al. 1968). Another hypothesis suggests that it is the result of a more general tendency in all nonstandard varieties of English to gravitate toward more primitive (i.e. not learned) linguistic patterns (Chambers 1995: 247). Still other hypotheses argue that *was/were* variation in some varieties is the result of innovative restructuring. For example in some dialects in the United States it has become differentiated according to polarity, with *was* used for positive contexts and *weren't* used for negative contexts, as in (2d) (Schilling-Estes and Wolfram 1994, Wolfram and Sellers forthcoming).

What is a viable way to evaluate these explanations? First, early research on this variability revealed that *was/were* variation occurred robustly in English vernaculars all over the world. However, as demonstrated in tables 28.2 and 28.3, overall distributions alone do not provide sufficient evidence to evaluate the mechanism that produces the variation from the underlying grammar. Second, *was/were* variation is the result of longitundinal linguistic variation and change in the English language as well. Thus, examination of the historical background is imperative in order to contextualize the synchronic situation. Moreover, detailed analysis modeling the internal grammatical constraints operating on the variation can reveal the extent and nature of the spread through linguistic structure (Labov 1982: 75). Again, comparative sociolinguistic analysis provides a means to disentangle these different hypotheses as well as, and perhaps more importantly, a means to constrain explanations.

## 7 Contextualizing Variation in Diachrony

At earlier stages in the history of English in Britain, variation in the forms of preterit indicative *be* was rampant (Curme 1977, Forsström 1948, Jespersen 1954, Pyles 1964, Visser 1963–73).

Use of *was* in 1st and 3rd person contexts was uniform across regions, but in 2nd person singular it varied according to geographic location. The southern regions of England mirrored contemporary standard English norms – the preterit indicative was *were* with all the plural personal pronouns (*we, you, they*) as well as 2nd person singular (*thou*). In the north and northeast, however, *was* was employed almost exclusively with 2nd person singular, as

in (3) (Forsström 1948). And this pattern continued into Early Modern English, as in (4).[4]

(3) Caym, Caym, thou *was* wode. (The Towneley Plays c.1450: 350)
(4) For in your last ye shew me that ye *was* troubled with ane swelling of the spleen. (Memorials of the family of Wemyss of Wemyss 1659)

Perhaps the most famous linguistic constraint on this variation, known as the Northern Subject Rule, involves a combination of the type of subject and the adjacency of the subject and verb to each other. According to Murray (1873: 211–12), "when the subject is a noun, adjective, interrogative or relative pronoun, or when the verb and subject are separated by a clause, the verb takes the termination *-s* in all persons." (see also Montgomery 1994, Tagliamonte 1999). Thus, *was* appeared after Full NPs as in (5), as opposed to plural pronouns, as in (6):

(5) The bernis both *wes* basit of the sicht. ("King Hart", Douglas, 1475–1522)
(6) They *wer* informed that my brother William his soun, should be a ward. (Letters on Duntreath, 1627)

While the association of *was* with 2nd person singular remained a northern feature, the tendency for plural NPs to have *-s* appears to have spread southward in the sixteenth and seventeenth centuries (see Godfrey and Tagliamonte 1999, Jones and Tagliamonte 2000).

A further constraint may be inferred from more recent studies. *Was/were* variation is sensitive to negation, but in contrasting ways depending on the dialect. Tagliamonte and Smith (2000) report a tendency for *was* to appear in negative, as opposed to affirmative, constructions in BCK, GYE, and NPR. Yet in contemporary non-mainstream varieties, both American (Schilling-Estes and Wolfram 1994) and British (Britain, forthcoming), a near categorical distinction between affirmative *was*, and negative *weren't* is reported. Indeed nonstandard use of *were* in negative contexts is reported in many British dialects (Hughes and Trudgill 1979: 63–5, Milroy and Milroy 1993).

Thus, three internal factors may be extrapolated from the literature.

1 *Was* was used almost exclusively in 2nd person singular in northern British dialects, while southern dialects used *were* (Brunner 1963, Forsström 1948, Mossé 1952).
2 *Was* occurred more often with plural NPs than pronouns. While originally a northern feature (Murray 1873), by the Early Modern English period it had spread throughout Britain (Visser 1963–73).
3 *Was* occurs in affirmative contexts; and *were* occurs in negative contexts, regardless of grammatical person in some dialects in the United States and Britain (Britain forthcoming, Schilling-Estes and Wolfram 1994).

Let us now test the data, once again using the comparative method in order to address broader questions regarding the nature of linguistic change, as well as processes of diffusion and/or innovation in time and space. The data sets available enable us to compare a number of dimensions: first, I compare two different geographic locations, North America and Britain. Second, I compare different varieties in North America according to their relative degree of participation in mainstream norms. Third, I compare according to African vs. British ancestry. Fourth, I contrast different dialect regions in Britain, north vs. south. Finally, I attempt to interpret these findings according to the putative origins of the major founding populations of the ancestors of the North American varieties, and thus in broad terms the composition of the original dialect input.

In the latter enterprise, I am abstracting away from the tremendous language and dialect contact of the early colonization period in the United States (e.g. Kurath 1928, 1949, 1964, Montgomery 1989a: 236) as well as the dialect mixture in Britain pre-dating the large scale migrations to North America in the 1700s. These, and other extralinguistic factors surely influenced the shaping of the varieties which emerged in North America in the early colonization period. However, to date there is no precise information about the nature of these factors as well as (and perhaps more importantly) any principled method for factoring their impact on linguistic structure into a quantitative analysis. Therefore, I focus on broad trends that can be inferred from migration patterns and population proportions from the historical record (e.g. Bailyn 1986, Bailyn and DeWolfe 1986) and rely on the details of the linguistic evidence for corroboration.[5]

## 8 Operationalizing Constraints on *was/were* Variation

First, consider the overall distribution of *was* in 2nd person singular and 1st, 2nd, and 3rd person plural in the five varieties targeted for investigation, figure 28.1. This view of the data reveals that *was/were* variation is robust across all the varieties under investigation. The relevant observation however is that the frequency of nonstandard use of *was* is not differentiated by broad geographic locale, i.e. North American vs. British, or even national locale, i.e. northern vs. southern Britain. Instead, the rates of nonstandard *was* are high everywhere, except in GYV, the variety which has evolved much more closely in tandem with mainstream developments in English than the other locales. This result suggests that the use of nonstandard *was* may be the result of differential contact with prescriptive norms. The overarching fact that use of nonstandard *was* is a pan-variety effect supports the hypothesis that *was/were* variation may be the result of vernacular primitives.

As with the analysis of *past temporal reference*, however, the overall distribution of forms tells us that all the varieties make extensive use of this nonstand-

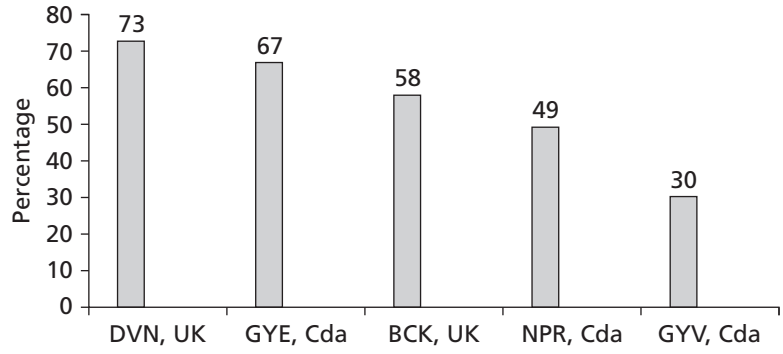


**Figure 28.1** Overall distribution of *was* in 2nd person singular, 1st, 2nd, and 3rd person plural

ard feature, but tell us little about possible internal constraints that differentiate the communities.

As discussed earlier, at least three linguistic features can be extrapolated from the historical and synchronic record. This presents three diagnostic constraints with which to compare and contrast varieties. First, consider the use of *was/were* according to grammatical person and number of the subject.

## *8.1 Grammatical person*

Figure 28.2 shows the distribution of *was,* but now separates the data into each of the grammatical persons for each of the communities. Where are the

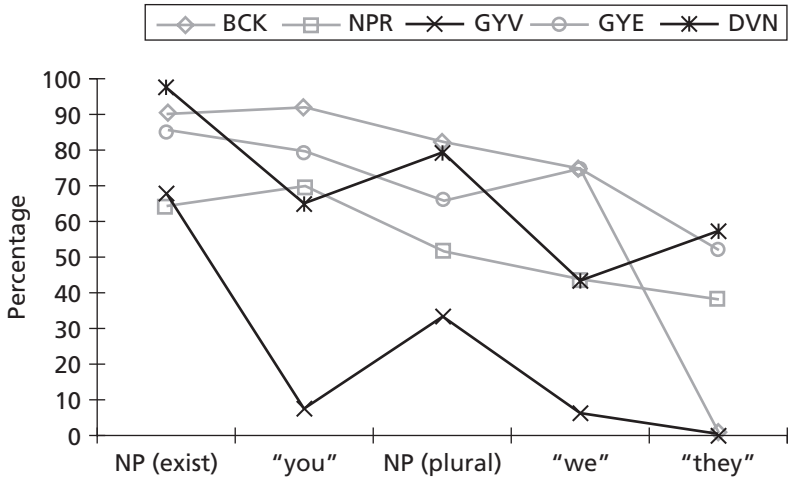


**Figure 28.2** Distribution of *was* by *grammatical person* of the subject

similarities and differences? These results reveal that some communities favor the use of *was* in certain contexts. First, consider plural existential constructions. This context favors the use of *was* across all communities, a correlation which has been traced back to the Old English period (Visser 1970), and which is widely reported in all contemporary varieties of English (Britain forthcoming, Britain and Sudbury 1999, Eisikovits 1991, Meechan and Foley 1994, Schilling-Estes and Wolfram 1994, Tagliamonte 1998b).

More importantly however, consider *non*-existential use of *was*. In three varieties, BCK, NPR, and GYE, this feature patterns according to a very similar hierarchy – 2nd person singular tends to have the highest rates, then 3rd person plural NPs, then 1st person plural, and finally 3rd person pronouns.[6] Most notably, *was* is highly favored in 2nd person singular. Moreover, it is this tendency which clearly sets these three communities off from GYV and DVN where 2nd person singular has one of the lowest rates of *was*. These results separate the varieties according to the differences between northern vs. southern British patterns, which in turn correspond to what we know about the general historical dialect roots of the ancestor populations of these communities.

Nevertheless, a single diagnostic provides only one piece of evidence. Consider another.

## *8.2 Type of subject*

Figure 28.3 displays the results for the distinction between full noun phrases and pronouns.

This historically-attested "Northern Subject Rule" is visible across all the communities. In every case, plural NP subjects exhibit a high, or relatively higher, frequency of use of *was* than with *they*.

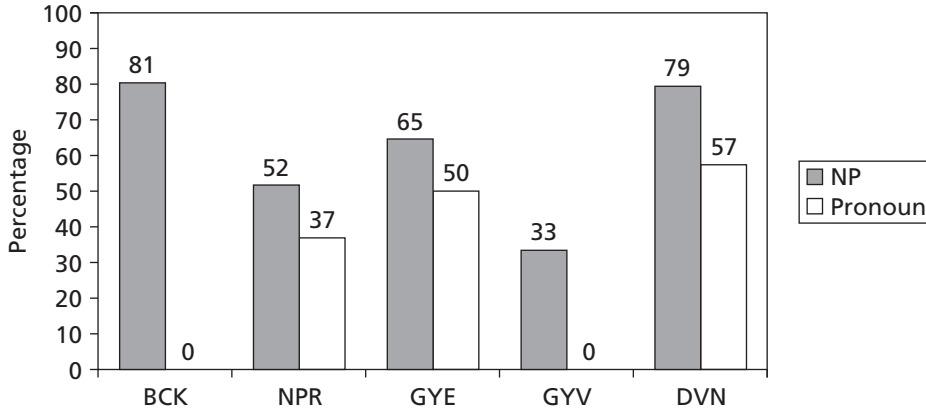


**Figure 28.3** Distribution of *was* by *full* NPs vs. 3rd person plural *pronouns*

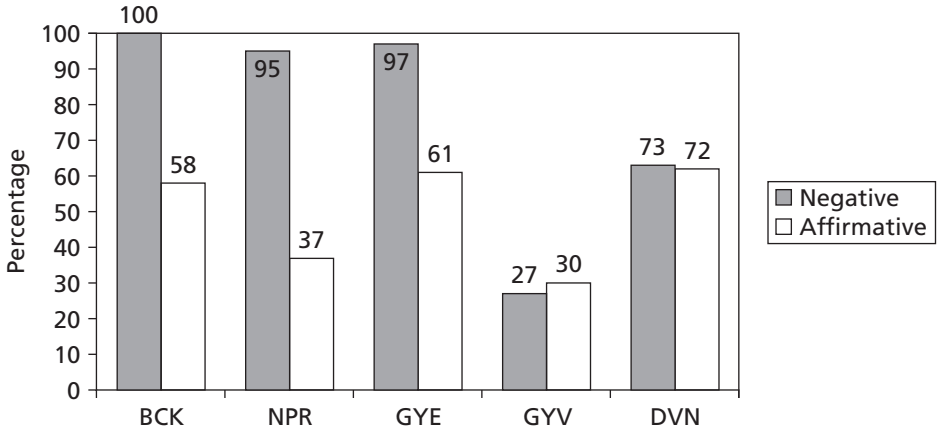


**Figure 28.4** Distribution of *was* by *negative* vs. *affirmative* contexts

BCK and GYV appear very similar in this feature in that they have categorical *were* with pronouns. However, plural NPs are the *only* location where *was* occurs in GYV (see figure 28.2). Whereas in BCK *was* is robust in every other context – a pattern that it shares with NPR and GYE. Thus, the relevant finding here is the *relative* high rates of *was* with NPs as opposed to *pronouns* precisely the same *direction* of effect attested in the historical record.[7]

### *8.3 The effect of negation*

Figure 28.4 displays the result for the distinction between negative and affirmative contexts. Once again, we observe a parallel pattern with BCK, NPR, and GYE: *was* is more frequent in negative, as opposed to affirmative contexts. In contrast, in GYV and DVN there is little to differentiate either affirmative or negative. None of the varieties exhibits the polarity effect found in Ocracoke or in southeast England.

## 9 Interpreting Similarities and Differences

The comparison of constraints on the use of *was* in all the varieties reveals that *was/were* variation is the result of systematic internal linguistic conditioning. It cannot be explained as the result of across-the-board regularization processes. The only pan-community effect is the strong propensity for *was* in existential contexts. Whether this is the result of diffusion, drift, or primitives remains an open question. However, in the rest of the verbal paradigm *was* exhibits a distinct and quite consistent set of constraints, and there are two sets of patterns

**Table 28.5** Variable *was*: Comparison of similarities and differences in internal linguistic features across communities

| | BCK | NPR | GYE | GYV | DVN |
|---|---|---|---|---|---|
| NP > Pronoun | ✓ | ✓ | ✓ | ✓ | ✓ |
| 2nd person singular | ✓ | ✓ | ✓ | ✗ | ✗ |
| Negation | ✓ | ✓ | ✓ | ✗ | ✗ |

which *contrast* across communities. Table 28.5 provides a checklist of which varieties share constraints and which do not.

Based on the parallel constraint rankings in NPR, GYE, and BCK Tagliamonte and Smith (2000) argued that the use of *was* in these varieties reflects the synchronic retention of linguistic patterns from northern varieties of British English, which would have embodied both the favoring effect of 2nd person singular *and* the favoring effect of plural NPs. The fact these communities also share the negation effect suggest that this too may have been one of the set of conditioning factors on *was/were* variation in these varieties, despite the fact that it had not been mentioned in the historical dialect literature.

However, such inter-variety parallelism does not unambiguously rule out the possibility that they represent universal constraints on variability (Wolfram 2000) or are the result of the strong vernacular tendencies discussed by Chambers (1995). Moreover, parallel constraints across one set of varieties is not conclusive proof of a common origin unless contrastive corroborating evidence can be found elsewhere. Thus, the additional data from Devon in the southwest bring crucial evidence to the comparative arena of *was/were* variation. How do the explanations offered by Tagliamonte and Smith (2000) fare with the Devon data added to the comparison?

First, figure 28.1 showed that this variety has extremely high rates of nonstandard *was* (73 percent), a rate which is on a par with the other non-mainstream varieties (BCK, NPR, and GYE), as well as many nonstandard North American dialects. However, the internal linguistic factors in figures 28.2 and 28.3 revealed that the ranking of constraints are not the same as the patterning reported for BCK, nor the African-American isolates in Nova Scotia (NPR and GYE).

All the varieties share the NP/PRO distinction, but with respect to the other two constraints, there is a clear demarcation between varieties. BCK, NPR, and GYE pattern together; while GYV and DVN pattern together. The question is why?

One way to interpret these linguistic results is to consider the history from which the varieties in these communities arose. While two (BCK and DVN) have remained in situ in Britain, the remainder are transported varieties whose original source may be traced back, at least in the first instance, to the United

States. The precise nature and proportion of the dialect features in contact in the colonial United States during the early formative period of American dialects is, at best, uncertain. Relevant to the present discussion, however, is the general fact that a disproportionate number of British northerners went to the early American backcountry or interior south (e.g. Bailyn 1986, Bailyn and DeWolfe 1986). As detailed in Poplack and Tagliamonte (2001), the original input settlers to NPR and GYE can be traced to the southern states. GYV, on the other hand, represents a variety which can be traced to the northern United States, an area which had a large proportion of input settlers from southern England.

The patterning of checks and crosses in table 28.5 reflects this. The three varieties which can be traced back to northern dialects of Britain or their input share the same three constraints, while GYV and DVN share a different set. This provides *linguistic* corroboration of the general sociohistorical facts and suggests a plausible interpretation of *was/were* variation in the North American context as being a retention of earlier dialect patterns traceable to the British Isles.

Thus, when internal linguistic constraints are consistently compared in data of appropriate size, character, and nature and if they also differentiate source varieties and/or regions, then they may hold the key to disentangling the thorny issue of dialect origins. Moreover, the convergence of evidence from three independent internal constraints which exhibit both parallel *and* constrastive patterns in a way which differentiates geographic dialects regions, diminishes the possibility that they have arisen by chance or are entirely the result of universals. Finally, such findings highlight how vital it is to consider linguistic features of English in terms of the highly differentiated regional dialects at earlier stages of the development of the language, back at its British source.

## 10 Using Factor Weights to Measure Grammatical Change

Up to this point we have been considering evidence from distributional analyses and constraints on morphosyntactic variation. Yet a great deal of variability in language comes from the fact that at any given point in the history of a language there are a tremendous number of linguistic features which are undergoing grammatical change. Such processes provide an invaluable area for cross-variety comparison.

Indeed, grammaticalization – a longitudinal process which may go on for centuries – necessarily produces variability in the grammar and this variability reflects the varying layers of grammaticalization attained by different forms (Hopper 1991: 23). Moreover, in the process of change, linguistic forms gradually

shift from one function to another. This trajectory can be viewed in the varying strength and distribution of independent linguistic features associated with one of the evolving grammatical morphemes. Indeed, such environmental correlations are held to be the keys to viewing the mechanism of diachronic grammaticalization in synchronic data (Traugott and Heine 1991). Moreover, because this process is gradual, sometimes lasting for centuries, the tracks of language change can be preserved across great distance and time. Practically speaking, such changes should be visible in an ordered series of shifts in factor weights (Labov 1982: 76). Thus, multivariate analysis which can model complex constraints and relative weights of numerous factors that operate simultaneously on linguistic features, provides an invaluable tool for actually tracking a grammaticalizing linguistic feature (see Labov 1972: 323).

## 11 Contextualizing Variation in Grammatical Change

Perhaps one of the best examples of ongoing linguistic change in English grammar is the *future reference* system. A number of different forms compete to mark contexts in which the speaker is making a prediction about an event which is yet to occur (Bybee and Pagliuca 1987), as in (7) (Poplack and Tagliamonte 1999, Tagliamonte 1997a).

(7) (a) She say, "if you looking for good you*'ll find* good . . . you looking for bad, you *gon find* bad. Ain't it true? (SAM/003/1282)
(b) It's like everything else. Some*'ll work*, and some *is* not *gonna* work. (NPR/074/1308–10)
(c) I think it*'s gonna get* worse before it*'ll get* better. (OTT/117/224B/17–20)
(d) I knew he *wasn't gonna be* any better, and he'd be an invalid all his life because I knew he *would never be* any- I thought I *was gonna be* sick right away. (GYV/101/B2A/7.07)
(e) I think she*'s gonna be* pretty cheeky. I think she*'ll be* cheeky. (YRK/O/04–23)

Although *will* and *shall* shared the *future temporal* reference system in English for centuries, the construction *going to* has more recently begun to encroach steadily on their functions. Interestingly enough, this change is not only well-documented, but a multitude of internal linguistic factors are implicated in its increasing frequency and gradual grammaticalization.

First, consider the diachronic picture. In early Old English future time was expressed by the simple non-past, with appropriate adverbial specification, for example:

(8) The ship sails *tomorrow*.

In this early period, two verbs *sceal* (shall) and *wille* (will) expressed present obligation and volition/willingness respectively. Gradually they both began to lose their intrinsic meaning (Traugott 1972, Visser 1963–73: 677) and entered into a long phase of variabilility. The two markers alternated according to the illocutionary act performed and the grammatical person of the subject noun (Lowth 1762/1967, Taglicht 1970, Zandvoort 1969).

This picture was complicated by the emergence of *going to*. First attested in the late 1400s, in future-in-the-past contexts, it became well-established with a wide array of lexical verbs by the seventeenth century while still retaining strong associations with its literal meaning of "intention" and "movement" (Danchev and Kytö 1994). However, *going to* as a future reference has been gaining ground ever since, particularly in the last century (Mair 1997).

## 12 Operationalizing Constraints on Grammaticalization

But precisely how *going to, shall*, and *will* compete as exponents of future expression has never quite been agreed upon. Factors contributing to different readings include: connotations of modality, degree of volition, certainty or prediction, intentionality, point of view, speaker attitude, probabilty or imminence of the event taking place, etc. Fortunately, at least some of these gradient semantic distinctions may also be observed in more mechanical subsystems of the grammar. I now demonstrate how such factors can be operationalized by focusing on syntactic and lexical features attested and/or claimed to have affected *future* expression (in general) throughout the history of English (see further detail Poplack and Tagliamonte 1999).

At least six constraints can be extracted from the literature. As mentioned earlier, the original association of *going to* with future-in-the-past is reflected in more contemporary observations. These suggest that *going to* used more frequently from the perspective of time passed than from the point of view of the present (Royster and Steadman 1923/1968), as in example (9a). Similarly, subordinate clauses are often discussed as being a favorable location for *going to*, perhaps in part due to the fact that many of the early attestations occurred in subordinate clauses, example (9b). As its meaning generalized from movement/intention to prediction, *going to* began to appear with non-human subjects, example (9c). This kind of lapse in restriction, here, on type of subject collocated with *going to*, is a common feature of grammaticalization, and shows up as the item is generalizing in meaning – subjects are no longer confined to animates capable of movement. Thus a propensity for *going to* to be favored with non-human subjects is a sign of further grammaticalization. Another factor heavily implicated in this grammatical change is the person and number of the subject.

Factors reflecting volition, or control, a reading said to be associated with *will*. First person subjects, as in example (9d), exercise it more, and so are predicted to occur more with *will*. Lexical content is another factor implicated in the grammaticalization of a form. Increased use of *going to*, originally a verb of motion, with another verb of motion, particularly *go* as in example (9e), implies bleaching or desemanticization of the original lexical meaning. Thus, a tendency toward such collocations is consistent with further grammaticalization. Finally, *going to* has long been associated with immediacy, example (9f).

(9) (a) They told him in the conference that uh, they *was going to give* him the bishop crown. (SAM/011/9014)
(b) I don't know whether she's *gointa teach* the sheep yoga or what! (YRK/t/20–04)
(c) They said his barn *was gonna burn* down. (/NPR/074/1244)
(d) I'm goin' up now and split now and I'*ll come* back and I'*ll get* a cup of tea or something or other and then I'*ll go back* up for another hour or so. (GYV/107/15.45)
(e) He's *gonna go* over there. (YRK/™/436,16)
(f) Now she's *gonna* make sandwiches and bologna. (GYE/048/404)

These detailed observations from the literature provide hypotheses about the development of *going to* which in turn provide critical diagnostics for the purpose of inter-variety comparison.

## 13 Language Change across Varieties

Given the linguistic pathways of this grammatical change described above, consistent comparison of the forms used for *future reference* in varieties which are distinguished by their relative degree of participation in mainstream norms may provide a view of this development. Such information can then be used to situate varieties vis-à-vis each other as well as to reveal important insights into the details of the mechanism of linguistic change.

Let us now consider the results when these factors are operationalized in a multivariate analysis. Once again, we compare across some of the same communities in table 28.1, this time the three enclaves: SAM in the Dominican Republic, and African Nova Scotian English as spoken in NPR, and GYE. These three varieties are expected to contrast with the rural variety spoken in GYV. Moreover, this time, we add to the comparison a control sample of Standard Canadian English as represented by a sample of elderly residents of Ottawa, Canada. This variety should reflect the most advanced stage of development of *going to*. The results are shown in table 28.6 (abstracted from Poplack and Tagliamonte 1999). In this table non-significant factors are in bold.

**Table 28.6** Five independent variable rule analyses of the contribution of factors selected as significant to the probability of *going to* in five North American varieties (factor groups selected as significant in bold)

| | Enclaves | | | Rural | Urban |
|---|---|---|---|---|---|
| | SAM | NPR | GYE | GYV | OTT |
| **Overall tendency:** | **0.59** | **0.55** | **0.50** | **0.31** | **0.48** |
| **Total *N*:** | **396** | **723** | **994** | **199** | **302** |
| **Point of reference** | | | | | |
| Past | **0.85** | **0.86** | **0.86** | **0.67** | **0.92** |
| Speech time | **0.44** | **0.40** | **0.43** | **0.45** | **0.40** |
| *Range* | *41* | *46* | *43* | *22* | *52* |
| **Type of clause** | | | | | |
| Subordinate | 0.58 | **0.68** | **0.69** | 0.59 | 0.55 |
| Main | 0.48 | **0.46** | **0.45** | 0.47 | 0.48 |
| *Range* | *10* | ***22*** | ***24*** | *12* | *7* |
| **Animacy of subject** | | | | | |
| Human | 0.50 | 0.50 | 0.50 | 0.50 | 0.48 |
| Non-human | 0.49 | 0.52 | 0.53 | 0.48 | 0.59 |
| *Range* | *1* | *2* | *3* | *2* | *11* |
| **Grammatical person** | | | | | |
| Non-first person | 0.50 | 0.51 | **0.57** | 0.54 | **0.61** |
| First person | 0.50 | 0.49 | **0.42** | 0.46 | **0.38** |
| *Range* | *0* | *2* | ***15*** | *8* | ***23*** |
| **Lexical content** | | | | | |
| Verb of motion | **0.34** | **0.33** | **0.35** | **0.32** | 0.51 |
| Other verb | **0.56** | **0.54** | **0.54** | **0.61** | 0.50 |
| *Range* | ***22*** | ***21*** | ***19*** | ***29*** | *1* |
| **Proximity in the future** | | | | | |
| Immediate | 0.47 | 0.49 | 0.53 | **0.66** | 0.59 |
| Non-immediate | 0.51 | 0.52 | 0.48 | **0.35** | 0.43 |
| *Range* | *4* | *3* | *5* | ***21*** | *16* |

*Source*: Poplack and Tagliamonte (1999)

I focus here on: (1) the constraint hierarchies of each of these effects, that is the ordering of factors within a linguistic factor group and (2) how this order reflects the direction predicted by the hypotheses in the literature. There are the noteworthy correspondences across varieties. Looking across the rows for each of the features under investigation, the relation of more to less in each factor (the constraint hierarchy), for each variety is virtually identical across the board.

There are only two places where these six varieties can be differentiated. The first is the effect of proximity in the future. Here, the five varieties partition into two groups, with the dividing line being between the three enclave varieties and the others. *Going to* is clearly favored for immediate future reference (consistent with prescriptive characterizations) in GYV and OTT; but there is no effect of temporal specialization in any of the enclaves. The second feature distinguishing the varieties is the effect of animacy. Only in OTT is *going to* favored for inanimate subjects, a context which is claimed to represent the most generalized and hence the most grammaticalized for *going to* (Bybee et al. 1994: 5).

The strength of these effects differs across varieties, as measured by the numbers for the *range*. *Point of reference* is one of the strongest constraints operating on the variation across all the varieties, while the effects of other factors shift in systematic ways. *Going to* rarely occurs with a (main) verb of motion. This avoidance of "redundancy," dating back to the time that *going to* was itself mainly perceived as a motion verb, is evidenced only in the enclave/rural communities. It has been neutralized in OTT. In contrast, an OTT innovation favoring *going to* with non-human subjects (.59) cannot be detected in any of the other varieties. These differences can be interpreted as a result of the fact that the different varieties are located at different points on the continuum of the grammaticalization of *going to.*

The position of GYV is pivotal. It shares its remoteness and relative isolation with the neighboring GYE (as well as with the other enclaves), but shares ethnic, racial, and other attendant characteristics with urban OTT. Interestingly, however, in its progress along the cline of grammaticalization, as measured by the range, Nova Scotian Vernacular English spoken in GYV appears to be more closely aligned with the African-origin enclaves: the effects of clause type and lexical content remain greater in these varieties when compared with OTT, while the effects of animacy and grammatical person of the subject have neutralized or are in the process thereof. On a fifth measure, point of reference, GYV has a much lower range than any of the others. Only on one measure, proximity in the future, is GYV aligned with urban OTT, i.e. along racial and ethnic lines.

These findings suggest that the language spoken by isolated speakers, whether of African or British origin, instantiates constraints that were operative at an earlier stage of the English language, and which are now receding from mainstream varieties.

Once again, we may ask if the comparative method provides conclusive evidence? As convincing as these inter-variety correspondences are, further exploitation of the comparative method can bolster the evidence even more.

The fact of the matter is that these findings are based on *separate* varieties of English spoken in different communities. These were in turn taken to reflect different points along the trajectory of grammaticalization of *going to* as a marker of *future reference* in the history of the English language. But how do we know that these inter-variety differences are actually the result of language change happening at different rates? Could they instead be the result of spontaneous parallel developments? Another interesting question is to what extent such claims can be corroborated by evidence from change in *apparent time* in *one* variety? As Bybee and Pagliuca (1987: 297) suggest, it is necessary to conduct an analysis of the use of grammaticalizing morphemes "as these changes are taking place." A further question is to what extent the conditioning factors reported in a North American context can be replicated on varieties of English elsewhere – particularly in the geographic context where the grammatical change originated.

## 14 Language Change in Apparent-time

In this section, I illustrate a quantitative analysis of the *future reference* system in a British variety of English (YRK, see table 28.1). Further, the data sample in table 28.7 was designed to include three broad age groups in order to examine the grammaticalization of *going to* in apparent-time in a single speech community.

In this demonstration the data set was coded, analyzed, and then configured for variable rule analysis to replicate the analyses in table 28.6. For the cross-generational comparison I focus on the relative *importance* of each factor (as indicated by its range), and above all, the extent to which its constraint hierarchy is shared by the other age groups. These findings will be interpreted in terms of the progress of each generation along the cline of grammaticalization of *going to* as a marker of *future reference*. The results are shown in table 28.8.

There are not only consistent parallels in the constraint ranking of factors across age groups, they are nearly identical to the hierarchy of constraints for the North American varieties in table 28.6. Moreover, these results are consistent with the known trajectory of *going to*. The same path is visible, in the

**Table 28.7** Distribution of sub-sample members

| | Male | Female | Total |
|---|---|---|---|
| 20–35 | 9 | 8 | 17 |
| 35–65 | 7 | 10 | 17 |
| 65+ | 10 | 10 | 20 |
| Total | 26 | 28 | 54 |

**Table 28.8** Three independent variable rule analyses of the contribution of factors selected as significant to the probability of *going to* in three age groups in the city of York (factor groups selected as significant in bold)

| | Britain (northern England) | | |
|---|---|---|---|
| | Young | Middle | Old |
| **Overall tendency:** | **0.36** | **0.32** | **0.25** |
| **Total *N*:** | **534** | **387** | **409** |
| **Point of reference** | | | |
| Past | **0.85** | **0.65** | **0.80** |
| Speech time | **0.45** | **0.48** | **0.44** |
| *Range* | *40* | *17* | *36* |
| **Type of clause** | | | |
| Subordinate | 0.56 | 0.55 | 0.52 |
| Main | 0.49 | 0.49 | 0.50 |
| *Range* | *7* | *6* | **2** |
| **Animacy of subject** | | | |
| Human | 0.50 | 0.56 | 0.51 |
| Non-human | 0.50 | 0.49 | 0.39 |
| *Range* | *0* | *7* | *12* |
| **Grammatical person** | | | |
| Non-first person | **0.58** | **0.58** | 0.52 |
| First person | **0.41** | **0.40** | 0.48 |
| *Range* | *17* | *18* | *4* |
| **Lexical content** | | | |
| Verb of motion | 0.53 | **0.29** | 0.47 |
| Other verb | 0.50 | **0.52** | 0.50 |
| *Range* | *3* | *23* | *3* |
| **Proximity in the future** | | | |
| Immediate | **0.61** | 0.48 | 0.46 |
| Non-immediate | **0.45** | 0.51 | 0.53 |
| *Range* | *16* | *3* | *7* |

*Source*: Tagliamonte (1997a)

behavior of each measure of grammaticalization in apparent-time just as it was visible across varieties.

Point of reference is one of the strongest effects on the choice of *going to.* The constraint ranking of type of clause is the same for all three age groups. However, it is the factors that are shifting in strength and significance that are most revealing. First, allocation of *going to* to proximate future reference is exhibited in the statistically significant effect amongst the youngest speakers only. Second, the neutrality between 1st vs. other grammatical subjects in the oldest generation shifts to a statistically significant, favoring effect of *going to* with 1st person subjects in the two younger generations. Third, the tendency for use of *going to* with animate nouns, apparent in the lower factor weights for non-human subjects in the oldest generations, is neutralized in the youngest generation. Such incremental alternations in apparent-time are consistent with the gradualness of grammatical change and reflect an ordered series of shifts in factor weights of the type noted by Labov (1982: 76). Moreover, each of the trends is comparable to the findings from the North American varieties studied previously. This provides additional corroboration that the effects are typical of English and part of the broader grammatical changes underway in the *future reference* system of the language.

Let us now compare the results from the British data and the North American data and focus on the factors most heavily implicated in grammatical change. Consider the factor weights for grammatical person. The patterning for the middle and youngest generations in YRK is similar to the elderly speakers in OTT, GYV, and GYE. Next, consider the oldest age group in York. Their patterning of constraints is quite distinct from the middle and younger age groups in the same community in that there is no effect. But notice how similar the overall pattern here is to SAM and NPR – two African-American isolates. Now, consider the proximity in the future results. Here, the youngest speakers in York are also not patterning with their elders. There is a distinct favoring effect of proximate future reference in the youngest generations, but not in the middle or oldest generation. Once again, the youngest generation in York are patterning like the elderly speakers from GYV and OTT, the two varieties representing a further advanced step along the grammaticalization cline for *going to*.

Thus, with respect to these two constraints which are highly implicated in the ongoing grammaticaliztion of *going to*, we can now make the observation that the *youngest* generation in Britain looks like the *oldest* generation in Canada. Perhaps even more surprisingly, however, is that the *oldest* generation in Britain looks much like the African-American enclaves in North America.

The two analyses in conjunction with one another provide corroborating evidence for a number of hypotheses. They suggest that grammatical change can be viewed in synchronic data. Further, the details of the lexical history of a grammaticalizing form appear to be reflected in variable constraints on its grammatical distribution. This may be viewed across sister varieties, as well as across different generations of the same community. The differences and

similarities across the generations in York *and* across varieties in both Britain and North America can be attributed to the fact that they reflect different points on the pathway of change of *going to* as a marker of *future reference* in English. This also lends support to the hypothesis that the relative *degree* of grammaticalization across *communities* may be related to the different ecological circumstances of their sociocultural history (Poplack and Tagliamonte 1999, Tagliamonte and Smith 2000). Finally, the multidimensional comparative perspective has revealed an additional and broader dimension. Grammaticalization of *going to* has progressed more quickly in North America than in Britain, and there appears to have been an acceleration of that change in the last 50–60 years. Further comparative research will undoubtedly fill in more of this emerging picture.

## 16 Conclusions

The studies I have summarized in this chapter apply a specific set of methodological principles to the study of language variation and change from a comparative cross-variety perspective. In any empirical discipline, hidden assumptions and details of method need to be laid out in an explicit way (see Lass 1993). The procedures discussed here are not new, but provide a detailed "unpacking" of the importance of accountability and proportional analysis and demonstrate the critical information provided by constraint ranking, the relative strength of effects and statistical significance. At the same time none of these lines of evidence is conclusive without additional procedures. Checking individual vs. group patterns, lexical effects, statistical fluctuation from small cells, interaction between constraints, and other problems are all a part of the method illustrated here, but I have not had enough space to elaborate on these details (see e.g. Poplack and Tagliamonte 2001). Any of these can mitigate confidence in the conclusions that can be drawn from a given data set. Thus, while a quantitative statistical method provides a powerful tool, any comparison is only as good as the accountability of the analysis that underlies it.

As language variation and change research develops further the application of quantitative methodology and a consistent comparative dimension will undoubtedly become more important, and techniques will become more refined. Such developments will increase the need for detailed linguistic criteria for determining the provenance or system membership of linguistic features. Moreover, as more data sets are discovered, collected and added to the body of materials available for analysis it will become even more critical to maintain rigorous and replicable standards in method. Although comparative reconstruction can be complicated by numerous factors, appropriate data exploited within a methodological framework such as described here can go a long way toward fulfilling the challenges of this continually evolving field.

## NOTES

1 The target of investigation will typically be varieties of a language, e.g. a dialect, but the comparison might also involve different age groups in a single community, different speakers interviewed at different points in time (e.g. Cukor-Avila 1997), or even different stages of acquisition (Hudson 1998).

2 Further information on the sociohistorical and linguistic characteristics of these varieties can be found in: Poplack and Sankoff 1987, Poplack and Tagliamonte 1989, 1991, 1994, 1999, 2001; Godfrey and Tagliamonte 1999, Jones and Tagliamonte 2000, Smith and Tagliamonte 1998, Tagliamonte 1998a, Tagliamonte and Smith 2000.

3 The issue of degree of isolation of SAM, NPR, and GYE is discussed more fully elsewhere (Poplack and Sankoff 1987, Poplack and Tagliamonte 1991, Poplack and Tagliamonte forthcoming). For further discussion of isolation in relation to BCK and DVN see (Godfrey and Tagliamonte 1999, Smith and Tagliamonte 1998) respectively.

4 Differentiation of 2nd person singular *was*, and 2nd person plural *were* is reported in a number of southern writers, but here its use was restricted and primarily stylistic (Petyt 1985, Pyles and Algeo 1993).

5 Support for a primary focus on linguistic evidence comes from a large-scale research project tracking the emergence of New Zealand English from diverse dialects which were in contact during the formative period. Trudgill et al. (2000) report that the mechanisms of dialect formation in the New Zealand context appear to have proceeded in a primarily deterministic fashion, regardless of external factors.

6 The exception is 1st person plural in GYE which has a heightened factor weight for *was* in comparison to BCK and NPR (for further discussion see Tagliamonte and Smith 2000)

7 Murray's (1873) observations acknowledge that the NP/Pro distinction is variable. His qualification is that *was* is present in NPs "though only as an alternative form."

## REFERENCES

Anttila, R. (1989). *An Introduction to Historical and Comparative Linguistics.* Amsterdam and New York: John Benjamins.

Bailey, G. and N. Maynor (1985). The present tense of *be* in Southern Black Folk Speech. *American Speech* 60, 3: 195–213.

Bailey, G., N. Maynor and P. Cukor-Avila (1989). Variation in subject-verb concord in Early Modern English. *Language Variation and Change*. 1, 3: 285–300.

Bailyn, B. (1986). *The Peopling of British North America.* New York: Alfred A. Knopf.

Bailyn, B. and B. DeWolfe (1986). *Voyagers to the West: A Passage in the Peopling of America on the Eve of the*

*Revolution*. New York: Alfred A. Knopf.
Bickerton, D. (1975). *Dynamics of a Creole System*. New York: Cambridge University Press.
Britain, D. (forthcoming). *Was/weren't* levelling in Fenland English. *Essex Research Reports in Linguistics*.
Britain, D. and A. Sudbury (1999). There's tapestries, there's photos and there's penguins: Variation in the verb BE in existential clauses in conversational New Zealand and Falkland Island English. Paper presented at Methods X. St John's, Newfoundland, Canada.
Brunner, K. (1963). *An Outline of Middle English Grammar*. Oxford: Basil Blackwell.
Bybee, J. L. and W. Pagliuca (1987). The evolution of future meaning. In A. G. Ramat, O. Carruba and G. Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam and Philadelphia: John Benjamins. 107–22.
Bybee, J. L., R. D. Perkins and W. Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
Chambers, J. K. (1995). *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Blackwell Publishers.
Clarke, S. (1997a). English verbal *-s* revisited: The evidence from Newfoundland. *American Speech* 72, 3: 227–59.
Clarke, S. (1997b). On establishing historical relationships between New World and Old World varieties: Habitual aspect and Newfoundland Vernacular English. In E. W. Schneider (ed.), *Englishes around the World I*. Amsterdam: John Benjamins. 277–93.
Clarke, S. (1997c). The search for origins: Habitual aspect and Newfoundland Vernacular English. Paper presented at NWAVE 26. Laval University, Quebec City, Canada.
Cukor-Avila, P. (1997). Change and stability in the use of verbal *-s* over time in AAVE. In E. W. Schneider (ed.), *Englishes around the World I*. Amsterdam: John Benjamins. 295–306.
Cukor-Avila, P. (1999). Reordering the constraints on copula absence in African American Vernacular English. Paper presented at Aix-en-Provence, France.
Curme, G. O. (1977). *A Grammar of the English Language*. Essex, Connecticut: Verbatim.
Danchev, A. and M. Kytö (1994). The construnction *be going to* + infinitive in Early Modern English. In D. Kastovsky (ed.), *Studies in Early Modern English*. Berlin: Mouton de Gruyter. 59–77.
Eisikovits, E. (1991). Variation in subject-verb agreement in Inner Sydney English. In J. Cheshire (ed.), *English around the World: Sociolinguistic Perspectives*. Cambridge: Cambridge University Press. 235–56.
Forsström, G. (1948). *The Verb "to be" in Middle English: A Survey of the Forms*. Lund: C. W. K. Gleerup.
Fries, C. C. (1940). *American English Grammar*. New York: Appleton, Century, Crofts.
Godfrey, E. and Tagliamonte, S. (1999). Another piece for the verbal *-s* story: Evidence from Devon in Southwest England. *Language Variation and Change*. 11, 1: 87–121.
Guy, G. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating Language in Time and Space*. New York: Academic Press. 1–36.
Guy, G. (1991). Explanation in variable phonology: An exponential model of

morphological constraints. *Language Variation and Change* 3, 1: 1–22.
Guy, G. and S. Boyd (1990). The development of a morphological class. *Language Variation and Change* 2, 1: 1–18.
Hazen, K. (1996). Dialect affinity and subject-verb concord: the Appalachian Outer Banks. *SECOL Review* 20, 1: 25–53.
Hickey, R. (ed.) (forthcoming). *The Legacy of English Dialects.* Cambridge: Cambridge University Press.
Hock, H. H. (1986). *Principles of Historical Linguistics*. Amsterdam: Mouton de Gruyter.
Hoenigswald, H. M. (1960). *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
Holm, J. (1975). Variability of the copula in Black English and its creole kin. *American Speech* 59, 4: 291–309.
Hopper, P. J. (1991). On some principles of grammaticization. In E. C. Traugott and B. Heine (eds.), *Approaches to Grammaticalization*, vol. 1: *Focus on Theoretical and Methodological Issues*. Amsterdam: John Benjamins. 17–35.
Hudson, R. (1998). Looking to the future: a study of the acquisition of Future Temporal Reference. MA Research Dissertation, University of York.
Hughes, A. and P. Trudgill (1979). *English Accents and Dialects: An Introduction to Social and Regional Varieties of British English.* London: Edward Arnold.
Jespersen, O. H. (1924). *The Philosophy of Grammar*. London: George Allen and Unwin.
Jespersen, O. H. (1954). *A Modern English Grammar on Historical Principles*, vol. 6: *Morphology.* London: George Allen and Unwin.
Jones, M. and S. Tagliamonte (2000). Comparative deconstruction: Disentangling the diagnostics of continuity and change. Paper presented at New Ways of Analyzing Variation (NWAV) 29, East Lansing, Michigan, USA. October 5–7.
Kurath, H. (1928). The origin of the dialect differences in spoken American English. *Modern Philology* 25: 385–95.
Kurath, H. (1949). *A Word Geography of the Eastern United States.* Ann Arbor: University of Michigan Press.
Kurath, H. (1964). British sources of selected features of American pronunciation: problems and methods. In D. Abercrombie and et al. (eds.), *In Honour of Daniel Jones: Papers Contributed on the Occasion of his Eightieth Birthday, 12 September 1961*. London: Longmans. 146–55.
Labov, W. (1966). *The Social Stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.
Labov, W. (1970). The study of language in its social context. *Studium Generale* 23, 1: 30–87.
Labov, W. (1972). *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.
Labov, W. (ed.) (1980). *Locating Language in Time and Space*. New York: Academic Press.
Labov, W. (1982). Building on empirical foundations. In W. Lehmann and Y. Malkiel (eds.), *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins. 17–92.
Labov, W., P. Cohen, C. Robins, et al. (1968). *A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City.* Final report, Co-operative Research Report 3288, Vol I. U.S. Regional Survey.
Lass, R. (1993). Projection: Procedures and constraints. In C. Jones (ed.), *Historical Linguistics: Problems and*

*Perspectives*. London and New York: Longman. 156–89.
Lowth, R. (1762/1967). *A Short Introduction to English Grammar*, no. 18. Menston, England: The Scholar Press.
Mair, C. (1997). The spread of the *going-to*-future in written English: A corpus-based investigation into language change in progress. In R. Hickey and S. Puppel (eds.), *Language History and Linguistic Modelling*. Berlin: Walter de Gruyter. 1537–43.
Meechan, M. and M. Foley (1994). On resolving disagreement: Linguistic theory and variation – *There's bridges*. *Language Variation and Change* 6, 1: 63–85.
Meechan, M., S. Tagliamonte and S. Poplack (1996). Isolating systems in language contact: A variationist method. Paper presented at Methods IX. Bangor, Wales.
Meillet, A. (1967). *The Comparative Method in Historical Linguistics*. Paris: Librairie Honoré Champion.
Milroy, J. and L. Milroy (1993). *Real English: The Grammar of English Dialects in the British Isles*. New York: Longman.
Milroy, L. (1980). *Language and Social Networks*. Baltimore: University Park Press.
Montgomery, M. B. (1989a). Exploring the roots of Appalachian English. *English World-Wide* 10: 227–78.
Montgomery, M. B. (1989b). The pace of change in Appalachian English. *English World-Wide* 10: 227–8.
Montgomery, M. B. (1994). The evolution of verbal concord in Scots. In A. Fenton and D. A. MacDonald (eds.), *Proceedings of the Third International Conference on the Languages of Scotland*. Edinburgh: Canongate Academic Press. 81–95.
Montgomery, M. B. (1997). Making transatlantic connections between varieties of English. *Journal of English Linguistics.* 25, 2: 122–41.
Mossé, F. (1952). *A Handbook of Middle English*. Baltimore: Johns Hopkins University Press.
Murray, J. A. H. (1873). *The Dialect of the Southern Counties of Scotland: Its Pronunciation, Grammar and Historical Relations*. London: Philological Society.
Neu, H. (1980). Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In W. Labov (ed.), *Locating Language in Time and Space*. New York: Academic Press. 37–54.
Patrick, P. L. (1999). *Urban Jamaican Creole: Variation in the Mesolect*. Amsterdam: John Benjamins.
Petyt, K. M. (1985). *Dialect and Accent in Industrial West Yorkshire*. Amsterdam: John Benjamins.
Poplack, S. (ed.) (2000). *The English History of African American English*. Oxford: Blackwell Publishers.
Poplack, S. and M. Meechan (1995). Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In L. Milroy and P. Muysken (eds.), *One Speaker, Two Languages*. Cambridge: Cambridge University Press. 199–232.
Poplack, S. and M. Meechan (eds.) (1998a). *Instant Loans, Easy Conditions: The Productivity of Bilingual Borrowing*, Special issue, *International Journal of Bilingualism*. London: Kingston Press.
Poplack, S. and M. Meechan (1998b). Introduction. How languages fit together in codemixing. Instant loans, easy conditions: The productivity of bilingual borrowing. *Journal of Bilingualism* 2, 2: 127–38.
Poplack, S. and D. Sankoff (1987). The Philadelphia story in the Spanish Caribbean. *American Speech* 62, 4: 291–314.

Poplack, S. and S. Tagliamonte (1989). There's no tense like the present: Verbal *-s* inflection in Early Black English. *Language Variation and Change* 1, 1: 47–84.

Poplack, S. and S. Tagliamonte (1991). African American English in the diaspora: The case of old-line Nova Scotians. *Language Variation and Change* 3, 3: 301–39.

Poplack, S. and S. Tagliamonte (1994). *-S* or nothing: Marking the plural in the African American diaspora. *American Speech* 69, 3: 227–59.

Poplack, S. and S. Tagliamonte (1999). The grammaticalization of *going to* in (African American) English. *Language Variation and Change* 11, 3: 315–42.

Poplack, S. and S. Tagliamonte (2001). *African American English in the Diaspora: Tense and Aspect*. Oxford: Blackwell Publishers.

Pousada, A. and S. Poplack (1982). No case for convergence: The Puerto Rican Spanish verb system in a language-contact situation. In J. A. Fishman and G. D. Keller (eds.), *Bilingual Education for Hispanic Students in the United States*. New York: Teachers' College, Columbia University. 207–37.

Pyles, T. (1964). *The Origins and Development of the English Language*. New York: Harcourt, Brace and World.

Pyles, T. and J. Algeo (1993). *The Origins and Development of the English Language*. Orlando: Harcourt Brace.

Rickford, J. (1986). Some principles for the study of Black and White speech in the south. In M. B. Montgomery and G. Bailey (eds.), *Language Variety in the South*. Alabama: University of Alabama Press. 38–62.

Rickford, J. R. (1991). Contemporary source of comparison as a critical window on the Afro-American linguistic past. In W. F. Edwards and D. Winford (eds.), *Verb Phrase Patterns in Black English and Creole*. Detroit: Wayne State University Press. 302–23.

Rickford, J. R. and R. Blake (1990). Copula contraction and absence in Barbadian English, Samaná English and Vernacular Black English. In K. Hall, J.-P. Koenig, M. Meecham, S. Reinman and L. A. Sutton (eds.), *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistic Society*. Berkeley: Berkeley Linguistics Society. 257–68.

Royster, J. F. and J. M. Steadman (1923/1968). The "going-to" future. In *Manly Anniversary Studies in Languages and Literature*. Freeport, NY: Books for Libraries Press. 394–403.

Sankoff, D. (1988). Sociolinguistics and syntactic variation. In F. J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*. Cambridge: Cambridge University Press. 140–61.

Santa Ana, Otto (1996). Sonority and syllable structure in Chicano English. *Language Variation and Change*. 8, 1: 63–90.

Schiffrin, D. (1996). Narrative as self-portrait: Sociolinguistic constructions of identity. *Language in Society* 25: 167–203.

Schilling-Estes, N. and W. Wolfram (1994). Convergent explanation and alternative regularization patterns: *Were/weren't* leveling in a vernacular English variety. *Language Variation and Change* 6, 3: 273–302.

Singler, J. V. (1991). Copula variation in Liberian Settler English and American Black English. In W. F. Edwards and D. Winford (eds.), *Verb Phrase Patterns in Black English and Creole*. Detroit: Wayne State University Press. 129–64.

Smith, J. and S. Tagliamonte (1998). "We were all thegither . . . I think we was all thegither": *was* regularization in

Buckie English. *World Englishes* 17: 105–26.
Tagliamonte, S. (1997a). Grammaticalization in apparent time: Tracing the Pathways of "gonna" in the city of York. Paper presented at Québec City, Canada.
Tagliamonte, S. (1997b). Obsolescence in the English Perfect? Evidence from Samaná English. *American Speech* 72(1): 33–68.
Tagliamonte, S. (1998a). Modelling an emergent grammar: Past temporal reference in St Kitts Creole in the 1780's. In P. Baker, A. Bruyn and N. Shrimpton (eds.), *St Kitts and the Atlantic Creoles*. Westminster: Westminster University Press. 201–36.
Tagliamonte, S. (1998b). *Was/were* variation across the generations: View from the city of York. *Language Variation and Change* 10(2): 153–91.
Tagliamonte, S. (1999). Back to the roots: What British dialects reveal about North American English. Paper presented at Methods X, St John's, Newfoundland, Canada. August 1–6, 1999.
Tagliamonte, S. and S. Poplack (1988). How Black English *past* got to the present: Evidence from Samaná. *Language in Society* 17(4): 513–533.
Tagliamonte, S., S. Poplack and E. Eze (1997). Pluralization patterns in Nigerian Pidgin English. *Journal of Pidgin and Creole Languages* 12(1): 103–29.
Tagliamonte, S. and J. Smith (2000). Old *was*; new ecology: Viewing English through the sociolinguistic filter. In S. Poplack (ed.), *The English History of African American English*. Oxford: Blackwell Publishers. 141–71.
Taglicht, J. (1970). The genesis of the conventional rules of "shall" and "will". *English Studies* 51(3): 193–213.
Thomason, S. G. and T. Kaufman (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley and Los Angeles, California: University of California Press.
Traugott, E. C. (1972). *A History of English Syntax: A Transformational Approach to the History of English Sentence Structures*. New York: Holt, Rinehart and Winston.
Traugott, E. C. and B. Heine (1991). Introduction. In E. C. Traugott and B. Heine (eds.), *Approaches to Grammaticalization*. Amsterdam: John Benjamins. 2–14.
Trudgill, P., E. Gordon, G. Lewis et al. (2000). Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36: 299–318.
Visser, F. (1963–73). *An Historical Syntax of the English Language*. Leiden: E. J. Brill.
Weinreich, U., W. Labov and M. Herzog (1968). Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin, TX: University of Texas Press.
Winford, D. (1992). Back to the past: The BEV/Creole connection revisited. *Language Variation and Change* 4: 311–57.
Wolfram, W. (1999). Principles of donor dialect attribution. Paper presented at Methods X, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.
Wolfram, W. (2000). Issues in reconstructing earlier African-American English. *World Englishes* 19(1): 39–58.
Wolfram, W. and D. Hatfield (1984). *Tense Marking in Second Language Learning: Patterns of Spoken and Written English in a Vietnamese Community*. Washington, DC: Center for Applied Linguistics.
Wolfram, W. and N. Schilling-Estes (to appear). Remnant dialects in the

coastal United States. In R. Hickey (ed.), *The Legacy of Colonial English: A Study of Transported Dialects*. Cambridge: Cambridge University Press.

Wolfram, W. and J. Sellers (1999). Ethnolinguistic marking of past be in Lumbee Vernacular English. *Journal of English Linguistics* 27: 94–114.

Wolfram, W., E. Thomas and E. Green (1997). Reconsidering the development of AAVE: Insights from isolated African American speakers. Paper presented at NWAVE 26, Québec City.

Young, R. and R. Bayley (1996). VARBRUL analysis for second language acquisition research. In R. Bayley and D. R. Preston (eds.), *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins. 253–306.

Zandvoort, R. W. (1969). *A Handbook of English Grammar*. London: Longman, Green.