Evaluation

Sociolinguists collect the speech they study from people engaged in ordinary activities, as we have seen in the previous section. The people may be reading aloud or writing, but they are more likely to be selling cars or discussing politics or gossiping. In a discipline that sets out to study language in its social uses, there really is no substitute for observing and recording ordinary speech.

Ordinary speech makes uncommon demands when it comes to analysis. For one thing, the tokens that carry social significance in speech events occur with wildly varying frequencies. A particular vowel phoneme may occur two or three times in every recorded minute, but a passive verb form may turn up only once or twice in an hour. For another thing, the social significance of these forms is very often an attribute not of their presence or absence in a person's speech but of their frequency in that speech compared to someone else's speech. For yet another, the differences that give speech its social significance are often minuscule. This is true not only of phonetic differences, in which it can be important to recognize that a vowel is slightly raised on one occurrence compared to another, or slightly more open. It is also true of grammatical differences, which are often carried by unstressed clitics that fade very quickly in the stream of speech.

Sociolinguists, like professional researchers in all empirical disciplines, develop refined sets of research skills that allow them to cope with the demands of their data. Most sociolinguists are exceptionally good at hearing vowel and consonant nuances in the speech stream, possibly because it is their aptitude for audio discrimination that attracts them to the work in the first place, but undoubtedly because they get exposed early and often to such nuances.

However, even the best ear can go no further in the analysis than assigning tokens to types. Much more than that is required, of course, and the study of actual speech has fostered a battery of analytic tools.

Observations involving social significance start with observations of frequencies, and frequencies require counting variants and correlating them with contextual features. Robert Bayley, in "The Quantitative Paradigm," discusses the principles that govern counting and correlating. He looks closely at VARBRUL, the logistic regression program expressly devised for handling variable data with distributional imbalances of the kind that inevitably accrue in real (as opposed to artificial) situations. VARBRUL has been developed

over some 30 years for sociolinguistic purposes, and along with it have come other quantitative methods for supplementing it and in some cases replacing it. Bayley discusses these too, and also some syntheses that are developing.

Variable frequencies are also at the heart of John R. Rickford's chapter on "Implicational Scales." The systematic nature of sociolinguistic data often reveals itself in robust implicational relationships between speech features, defined schematically by the formula "If X, then Y but not vice-versa." So, in Rickford's focal case study, Jamaican Creole speakers who use *no ben* for "didn't" inevitably use *pikni* for "child", but there are other speakers who use *pikni* and never use *no ben*. Implicational relationships like these constrain the range of variation that actually occurs in a dialect continuum, and reveal the structure in variability.

Erik R. Thomas in "Instrumental Phonetics" discusses the identification of sociolinguistic tokens through the use of acoustic technology of several kinds. He reviews the application of instrumental methods in vowel variation, a fairly well-developed area, but advocates and explores additional acoustic applications that could be imported into sociolinguistics or implemented more extensively. Although some sociolinguists have exploited the rapid developments that have taken place in instrumental analysis, the possible uses, as Thomas points out, are considerably broader, and the potential gains in accountability and accuracy are inestimable.

J. K. Chambers

5 The Quantitative Paradigm

ROBERT BAYLEY

The quantitative paradigm in sociolinguistics originated in the studies conducted by William Labov in New York and Philadelphia in the 1960s and 1970s (Labov 1966, 1969a, 1972a, 1972b). This approach to the study of language was subsequently extended to a wide variety of language communities around the world, including Panama (Cedergren 1973), Norwich, England (Trudgill 1974), Anniston, Alabama (Feagin 1979), Guyana (Rickford 1987), and Rio de Janeiro (Guy 1981), to name just a few. The central ideas of this approach are that an understanding of language requires an understanding of variable as well as categorical processes and that the variation that we witness at all levels of language is not random. Rather, linguistic variation is characterized by orderly or "structured heterogeneity" (Weinreich et al. 1968: 99–100). That is, speakers' choices between variable linguistic forms are systematically constrained by multiple linguistic and social factors that reflect underlying grammatical systems and that both reflect and partially constitute the social organization of the communities to which users of the language belong. In addition, synchronic variation is often a reflection of diachronic change (Labov 1994).

In this chapter, I outline the assumptions underlying this approach to the study of language variation and change. I then focus on methods of quantitative analysis, with an emphasis on variable rule, or VARBRUL, analysis, the most common method of multivariate analysis in quantitative sociolinguistics. The next section considers alternative methods that have been recently proposed to overcome some of the limitations of VARBRUL. Finally, I examine recent work that synthesizes traditional approaches to the study of linguistic variation and ethnography.

1 Theoretical Principles of the Quantitative Paradigm

Several key principles underlie the quantitative study of linguistic variation. Among the more important are the "principle of quantitative modeling" and

the "principle of multiple causes" (Young and Bayley 1996: 253). The "principle of quantitative modeling" means that we can examine closely the forms that a linguistic variable takes, and note what features of the context co-occur with these forms. By context is meant the surrounding linguistic environment and the social phenomena that co-occur with a given variable form. With a large enough set of data, we are able to make statements about the likelihood of co-occurrence of a variable form and any one of the contextual features in which we are interested.

These statements express in quantitative terms the strength of association between a contextual feature and the linguistic variable. For example, Bayley and Pease-Alvarez (1997), in a study of Mexican immigrant and Chicano Spanish, were interested in the relationship between the degree of discourse connectedness, operationalized as continuity of subject, tense, and mood with the preceding tensed verb, and the likelihood that Mexican-born and Chicano children would use an overt pronoun rather than a null subject in sentences such as (1) and (2):

(1) *una noche cerca de navidad ella/Ø nos dijo que se sentía muy mal . . .*
one night near Christmas she told us that she felt very bad . . .

(2) *entonces él/Ø tuvo que cerrar la ventana.*
then he had to close the window.

After conducting a VARBRUL analysis, Bayley and Pease-Alvarez reported that in cases where there was continuity of subject, tense, and mood, the weight of the factor was .244. On the other hand, in cases with there was a change in discourse topic, the likelihood of an overt pronoun was .653. Factor weights favor the variant when they exceed .50 and otherwise disfavor it (as discussed in more detail below). This means that an overt pronoun is highly unlikely to occur when continuity of subject, tense, and mood is preserved. However, overt pronouns are quite likely to occur when the discourse topic changes. Moreover, as shown in table 5.1, in the narrative discourse that Bayley and Pease-Alvarez investigated, the likelihood of an overt pronoun increases as the degree of discourse connectedness decreases. To the extent that these results are representative of other Mexican-born and Chicano children, we may expect that the use of overt pronouns in the speech of these children will pattern in the same way.

The second principle, the "principle of multiple causes," means that it is unlikely that any single contextual factor can explain the variability observed in natural language data. For example, in Bayley and Pease-Alvarez's study, the degree of discourse connectedness with the preceding tensed verb was not the only significant constraint on variation between null and overt pronouns. Person and number, verb type (present; preterit; imperfect, conditional, or subjunctive), immigrant generation, and speaker gender also proved to have statistically significant effects.

Table 5.1 Degree of discourse connectedness and overt pronoun use in Mexican immigrant and Chicano children's Spanish

Degree of discourse connectedness	% overt pronoun	VARBRUL weight
First degree: continuity of subject, tense, and mood	12	.293
Second degree: continuity of subject, different tense and/or mood	20	.405
Third degree: subject continuity interrupted by one or more intervening clauses	21	.490
Fourth degree: last occurrence of subject in another syntactic function	35	.607
Fifth degree: change in narrative section or discourse topic	32	.653
Total/input probability	24	.198

Source: Bayley and Pease-Alvarez (1997: 360)

The great majority of studies of linguistic variation has shown that the variables that have been closely examined, like null pronoun variation in Spanish, are subject to not one, but many contextual conditioning factors. For example, studies of a variety of English dialects have shown that final consonant cluster reduction, or -t,d deletion, is subject to a wide range of linguistic factors that exhibit remarkable cross-dialectal consistency (see e.g. Bayley 1994a, Guy 1980, 1991, 1997, Labov 1989, 1997, Labov et al. 1968, Roberts 1997, Santa Ana 1992, Wolfram 1969, Wolfram and Fasold 1974). In most dialects in which this variable has been studied, -t,d is far more likely to be deleted if it is part of a monomorpheme, as in *mist*, than if it is a past tense marker, as in *missed*. Deletion is also subject to phonological constraints. Final -t,d is more likely to be deleted if the following segment is a consonant than if it is a vowel. Other linguistic factors also influence the likelihood of -t,d deletion. The complexity of the multiple factors is succinctly illustrated by Labov's summary of the pan-English pattern, using variable rule notation (1989: 92):[1]

/-t,d/ → <Ø>/<str.> (C) <–cont. +cons.> <cat.> _ <features>

<α voi> <α voi>

a *b* *c* *d* *e*

f *f*

a *syllable stress* (unstressed > stressed)
b *cluster length* (CCC > CC)

c the phonetic features of the *preceding consonant*, yielding the segmental order /s/ > stops > nasals > other fricatives > liquids
d the *grammatical status* of the final /-t,d/, with the order: part of *-n't* morpheme > part of stem > derivational suffix > past tense or past participial suffix
e the phonetic features of the *following segment*, yielding the order: obstruents > liquids > glides > vowels > pauses
f *agreement in voicing* of the segments preceding and following the /-t,d/ (homovoiced > heterovoiced)

Not only is deletion constrained by the grammatical status of final -t,d and the features of the following segment, it is also constrained by syllable stress, with -t,d in unstressed syllables more liable to deletion than -t,d in stressed syllables, and by cluster length, with triclusters more liable to reduction than biclusters. In addition, -t,d deletion is affected by the phonetic features of the preceding segment and by voicing agreement of the segments preceding and following the variable.

In addition to the principles of quantitative modeling and multiple causes, two other principles are critical to the variationist paradigm. These are summarized by Guy (1991):

- Individual speakers may differ in their basic rate of use of a variable rule, that is, in their input probability for the rule.
- Individuals should be similar or identical in the factor values assigned to linguistic constraints on the rule. (This assumption is usually qualified to apply just to people who belong to the same speech community.)

The first of these principles offers a way to understand how groups of speakers who use a particular variant at very different rates may be regarded as members of the same speech community. For example, Wolfram found that Detroit African-Americans deleted -t,d at very different rates, depending on the social class to which they belonged, as well as on a number of linguistic factors. Table 5.2 shows his results for social class, grammatical function, and following phonological environment, expressed in percentages. Figure 5.1 presents the same information graphically.

In these results, we can see that the rate of -t,d deletion by these speakers is affected by all three of the factors examined. When -t,d is a past tense morpheme and is followed by a vowel, the deletion rate for upper-middle-class (UM) speakers is only 7 percent. In contrast, lower-working-class (LW) speakers delete -t,d in the same environment at a 34 percent rate. When -t,d is not a past tense morpheme and the following segment is a consonant, the rate of deletion increases to 79 percent for upper-middle-class speakers and to a near categorical 97 percent for lower-working-class speakers. Note, however, that the linguistic factors have the same effect on speakers of all social classes, despite differences in the overall percentages of deletion. Regardless of social class, the order of environments for final -t,d deletion is: –past, +cons > –past,

Table 5.2 Percentages of -t,d deletion in Detroit African-American English by linguistic environment and social class

Environments	Social class			
	Upper-middle	Lower-middle	Upper-working	Lower-working
Following vowel:				
-t,d is past morpheme (e.g. "missed in")	7	13	24	34
-t,d is not past morpheme (e.g. "mist in")	28	43	65	72
Following consonant:				
-t,d is past morpheme (e.g. "missed by")	49	62	73	76
-t,d is not past morpheme (e.g. "mist by")	79	87	94	97

Source: Wolfram and Fasold (1974: 132)

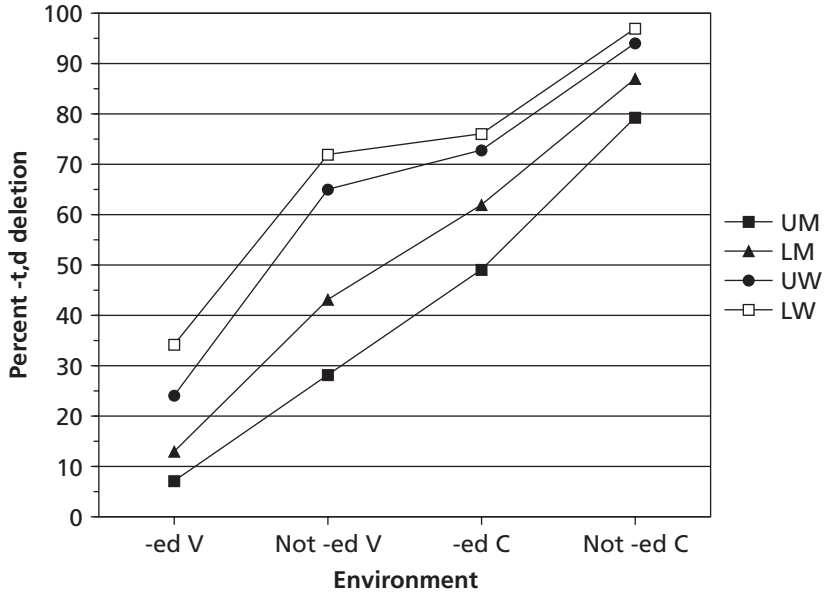

Figure 5.1 -t,d deletion in Detroit African-American English by linguistic environment and social class

Source: Wolfram & Fasold, 1974: 132

+vowel > +past, +cons > +past, +vowel. With respect to this aspect of the grammar, then, members of different social classes can be said to belong to the same speech community, even though they differ considerably in their overall rate of -t,d deletion. The speakers Wolfram studied thus provide a clear example of Labov's definition of a speech community:

> The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms: these norms may be observed in overt types of evaluative behavior, and by the uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage. (Labov 1972b: 120–1)

Thus far, the discussion has concerned social groups rather than individuals. However, it is conceivable that the percentages in studies such as Wolfram's might be arrived at by averaging speakers who happened to belong to the same social group but who exhibited very different behavior with respect to particular linguistic variables. For example, the -t,d deletion rate of 43 percent by lower-middle-class speakers when the -t,d is not a past tense morpheme and the following segment is a vowel might be arrived at by combining tokens from speakers who delete -t,d at a rate of 86 percent with an equal number of tokens from speakers who never delete -t,d in this environment. In practice, however, this does not happen. Guy (1980), for example, examined -t,d deletion in a sample of New Yorkers and Philadelphians. His results showed that, as long as a sufficient number of tokens was available (approximately 20 per cell), results for individuals closely matched the group pattern. Later studies, including my own work on the role of grammatical aspect in the acquisition of past-tense marking by Chinese learners of English (Bayley 1994b), have confirmed Guy's finding concerning the relationship between group and individual patterns of variation.

To summarize, research in the quantitative paradigm has demonstrated the systematic nature of much of the linguistic variation that was previously thought to be random. Moreover, research has shown in fine detail that variable linguistic forms are constrained by multiple internal and external factors. And, research has shown that, at least with respect to major linguistic constraints, given sufficient data, individual patterns do in fact match group patterns. These insights have been gained by adopting certain methods of analysis. Perhaps most obvious of these is the focus on actual language as produced by speakers in communities rather than on linguistic intuitions, or grammaticality judgments, as has been the practice in formal linguistics. The data gathered through sociolinguistic interviews of the type pioneered by Labov (1963, 1966), the social network approach developed by James and Leslie Milroy in Belfast (Milroy 1987), or the intense participant observation exemplified by Eckert's (2000) study of a Detroit area high school have been subject to a variety of interpretations based on different theoretical principles. In its concentration on language as used by members of the communities under study, however,

research in the quantitative paradigm has remained resolutely "secular," to use Labov's term. That is, regardless of the theoretical predisposition of the researcher, work in the quantitative tradition has tended to preserve the principle of accountability (Sankoff 1990), which involves dealing with the full range of variability present in the (relatively) informal interactions of language users. Given the multitude of factors, both linguistic and social, that can potentially influence a language user's choice of one or another variable form, adherence to the principle of accountability necessitates multivariate analysis.

2 Quantitative Analysis

The quantitative modeling of the correlations between language variation and the multiple contextual factors that promote or inhibit use of a particular variant is no easy matter. In studies that relate variation to a single contextual factor, a simple statistical procedure such as a comparison of two means with the help of a t-test has been used (e.g. by Beebe 1977). However, such a model is inadequate when multiple influences are likely to be involved. Analysis of variance is another technique that has been used (e.g. Tarone 1985) to relate variation to a single independent variable with multiple levels. In principle, it is possible to extend an analysis of variance to additional variables, but with the kind of data usually collected in studies of linguistic variation, this is hardly ever practicable. An example should help make clear why this is the case. In a study of -t,d deletion in Tejano/Chicano English (Bayley 1994a), I originally hypothesized that the variation would be influenced by eleven separate independent variables, each of which had theoretical and empirical support from previous studies. The eleven independent variables were all nominal (that is to say they could be further subdivided into two or more categories) and were as follows:

Morphological class: monomorpheme, semiweak verb (e.g. *left*), past tense or past participle, *-n't*;
Phonetic features of the preceding segment: /s/, nasal, stop, fricative, /r/, /l/;
Phonetic features of the following segment: consonant, /l/, /r/, glide, vowel, pause;
Syllable stress: unstressed, stressed;
Voicing agreement of the preceding and following segments: homovoicing, heterovoicing;
Cluster length: CCC, CC;
Speech style: conversation, reading continuous passage, word list;
Reported first language: English, Spanish;
Current home language: English, English and Spanish, Spanish;
Gender: male, female.
Age: 14–24, over 25.

In this model there are 11 separate factor groups (independent variables) comprising a total of 34 separate factors (categories). The number of possible combinations of factors (also known as cells) is 82,944. This is an extremely large number of cells for a multiple ANOVA to handle. In addition, most cells are empty, although nearly 5,000 tokens of the dependent variable – final consonant clusters – were collected for the study. This is because many combinations are linguistically impossible or highly unlikely, leaving more than 80,000 cells with missing data. Moreover, the majority of the filled cells represent only one token of the dependent variable, presence or absence of final -t,d. Algorithms for calculating ANOVA normally require equal numbers of tokens in each cell and are clearly inapplicable to such a case. Even algorithms for calculating unbalanced ANOVAs will fail when faced with such extreme distributional imbalances. ANOVA is a statistical procedure designed to deal with the kind of balanced data that emerge from controlled experiments. It is inadequate to handle the kind of naturally occurring data that are collected in studies of sociolinguistic variation.

3 Multivariate Analysis with VARBRUL

Modeling linguistic variation can be carried out by a number of commercial statistical software packages, usually under the name of logistic regression (e.g. Norušis and SPSS 1996, SAS Institute 1996). However, the programs known as VARBRUL have been used most extensively in sociolinguistics because they have been deliberately designed to handle the kind of data obtained in studies of variation. They also provide heuristic tools that allow the investigator to modify hypotheses and reanalyze the data easily. The statistical bases for the VARBRUL programs are set out in Sankoff (1988), and the procedures for using the software are explained in detail in Young and Bayley (1996) and in the documentation that accompanies the programs. The two most widely available versions are GoldVarb for the Macintosh (Rand and Sankoff 1991) and VARBRUL for the PC (Pintzuk 1988).[2]

A full explanation of the steps involved in carrying out a multivariate analysis with VARBRUL is beyond the scope of this chapter. Here, the discussion will be limited to addressing several questions that arise in any study, including defining the envelope of variation, testing for significance, interpreting the results, and dealing with the limitations inherent in the program. Readers who wish to pursue the topic in greater depth should consult the extensive literature on variable rule analysis and use of the VARBRUL programs (e.g. Cedergren and Sankoff 1974, Guy 1980, 1988, 1993, Rousseau 1989, Rousseau and Sankoff 1978, Sankoff 1988, Sankoff and Labov 1979, Young and Bayley 1996).

The first steps in conducting a VARBRUL analysis are to define the variable and the envelope of variation. That is, what forms count as instances of the variable? Are the forms that vary indeed two ways of saying the same thing?

In many studies, particularly studies of phonological variation, defining the envelope of variation is not a problem. For example, *fishing* and *fishin'* clearly have the same referential meaning, as do *west side* and *wes' side*. However, it becomes less obvious that variable forms meet the criterion of being two ways of saying the same thing at higher levels of linguistic structure. VARBRUL has been used to analyze variation in syntax, discourse, and code-switching (e.g. Poplack 1980, Poplack and Budzhak-Jones 1998, Schiffrin 1982, Weiner and Labov 1983). However, the use of VARBRUL analysis for modeling variation in syntax in particular has given rise to considerable controversy (see e.g. Labov 1978, Lavendera 1978). In fact, the problem of defining what counts as an instance of the variable may affect the study of even such frequently examined variables as copula deletion in African American Vernacular English (AAVE), which has been referred to as a "showcase variable of American dialectology and quantitative sociolinguistics" (Rickford et al. 1991: 104).

Rickford et al. (1991) examined the quantitative consequences of different decisions about defining the envelope of variation for this often studied variable, as well as underlying models, i.e. whether variable AAVE copula absence results from a deletion or an insertion rule[3] and whether *is* and *are* should be analyzed separately or combined. Rickford et al. (1991) performed nine separate multivariate analyses of 1,424 tokens extracted from interviews and peer group sessions with approximately 30 speakers in East Palo Alto, California. Their results for *is* and *are* showed that these two variants could best be accounted for within a single model that included a person-number factor group to account for the different forms. However, to perform the analyses, they found it necessary to exclude approximately 2,000 tokens in addition to nonfinite and past tense forms of *be*. Although Rickford et al. (1991) recognized that excluded tokens were relevant to the question of whether AAVE has an underlying copula, it was nevertheless necessary to exclude them from the quantitative analyses because they showed invariant copula presence (e.g. *am* was present in contracted form nearly 100 percent of the time) or because it was impossible to determine whether the copula was present or not (e.g. cases of contracted *is* followed by a sibilant) (1991: 107).

The second issue that arises early in a study concerns specifying the factors that may potentially influence the choice of a variant. In general, it is best to be liberal at this stage, although each factor group should be based on a well-motivated hypothesis. Lucas (1995), for example, investigated the potential effects of eight separate factor groups on the choice of a variant of the sign DEAF in American Sign Language. As it turned out, most of these groups proved not to be statistically significant. However, the labor of coding for many factors was not expended in vain. The study demonstrated that Liddell and Johnson's (1989) claim that variation in the form of DEAF is influenced primarily by the location of the preceding sign, a phonological constraint, is at best incomplete. Lucas also demonstrated the previously unsuspected influence on the choice of variant of the grammatical category to which DEAF belongs,

a finding that was later confirmed in a larger study based on a representative sample of the Deaf community in the United States (Bayley et al. 2000).

Once coding is complete and the data are entered into the program, VARBRUL estimates the factor values (or probabilities) for each contextual factor specified (e.g. the phonetic features of the following environment or the social class to which a speaker belongs). This is done by combining the input probability (p_0, the likelihood that the "rule" will apply regardless of the presence or absence of any other factor in the environment) with the applicable factor weights from each of the factors in the model ($p_1, p_2, \ldots, p_n$), according to the formula:

$$p = \frac{p_0 \times \ldots \times p_n}{[p_0 \times \ldots \times p_n] + [(1 - p_0) \times \ldots \times (1 - p_n)]}$$

The program provides a numerical measure of the strength or influence of each factor, relative to other factors in the same group, on the occurrence of the linguistic variable under investigation. Values range between 0 and 1.00. A value, or weight, between .50 and 1.00 indicates that the factor favors use of a variant relative to other factors in the same group. For example, Baugh (1983) examined -t,d deletion in AAVE, among other variables. Among the factor groups for which he coded were the grammatical function of the word containing the -t,d cluster and the type of speech event from which the data were extracted. The factors in the grammatical function group were monomorphemes, e.g. *mist*, *past*, semiweak verbs, e.g. *kept*, *lost*, and past tense forms, e.g. *missed*, *passed*. Baugh divided the speech events into four types, depending on the speakers' familiarity with one another and the extent to which they participated in African-American vernacular culture. He hypothesized that participants in Type 1 events, characterized by familiarity of the speakers and shared participation in African-American vernacular culture, would favor use of vernacular forms, in this case, -t,d deletion. Conversely, he hypothesized that vernacular forms would be less likely to occur in Type 4 events, where the speakers were not well-acquainted and where AAVE was not common to all. Results for these two factor groups are shown in table 5.3.

Baugh reported that the results for the grammatical function group were significant. Like speakers of other English dialects, speakers of AAVE are more likely to delete final -t,d when it does not carry any grammatical meaning, as is the case in monomorphemic words. They are less likely to delete -t,d when it functions as a past tense ending. Semi-weak, or ambiguous, verbs, which are characterized by an internal vowel change and affixation of -t,d, have an intermediate value. The type of speech event, however, failed to reach significance. The values for all four types of speech events hover around .5. Contrary to Baugh's hypothesis, -t,d deletion was *not* affected by this factor.

In addition to calculating values or weights for each factor, VARBRUL also calculates the input probability, which, as noted above, is the overall likelihood that speakers will choose the variant selected as the application value (the value

Table 5.3 -t,d deletion by grammatical function and speech event type in African-American Vernacular English

Factor group	Factor	VARBRUL weight
Grammatical function	No grammatical function, e.g. *past*	0.683
	Ambiguous function, e.g. *lost*	0.523
	Past tense function, e.g. *passed*	0.353
Speech event type	**Type 1**: Familiar participants, all of whom are natives of African-American vernacular culture	0.482
	Type 2: Participants are not well acquainted, but all are members of African-American vernacular culture	0.523
	Type 3: Participants are well acquainted, but do not share AAVE	0.499
	Type 4: Participants are not well acquainted and AAVE is not common to all	0.496

Source: Baugh (1983: 98)

that counts as an application of the "rule" being investigated). In my own study of -t,d deletion in Tejano English, for example, the input probability was .469, indicating that -t,d would be likely to be deleted nearly half the time regardless of the presence or absence of any other factor in the environment.

The program also provides several measures of goodness of fit between the model and the data. These include the total chi-square, the chi-square per cell, and the log likelihood. The total chi-square measures the degree of interaction among factors from different factor groups. An acceptable value for the total chi-square is derived by looking at a table of the chi-square distribution at the desired probability level (say $p < .05$) and the appropriate number of degrees of freedom in the model. The degrees of freedom in any VARBRUL model are calculated by subtracting the number of factor groups from the total number of factors. For example, in the case of a model with 6 factors distributed among 3 binary factor groups, the number of degrees of freedom is 3. From a chi-square table, we can see that the total chi-square should be less than 7.815 for us to accept that our model has less than 1 chance in 20 of being right by chance ($p < .05$).

The chi-square per cell figure is simply calculated by dividing the total chi-square by the number of cells. The lower the chi-square per cell figure, the less likely there is interaction among factors. As a general rule, a chi-square per

cell figure greater than 1.5 suggests that there may be an interaction between two or more factor groups, e.g. ethnicity and social class. In such cases, it may be necessary to recode the data to remove the interaction. For example, rather than have separate factor groups for ethnicity and social class, it may be better to combine them. Thus two binary factor groups for African-American and Euro-American and for working and middle class speakers might be combined to form a single factor group consisting of African-American working class, African-American middle class, Euro-American working class, and Euro-American middle class.[4] Finally, VARBRUL provides the log likelihood statistic, also a measure of goodness-of-fit. Figures closer to zero represent better models than log likelihoods further removed from zero (see Young and Bayley 1996: 272–3).

The factor values and input probability reported in a VARBRUL run provide useful information. They are not sufficient, however, to confirm or disconfirm the hypotheses that led to the inclusion of the factors in the original model of variation. Our goal in VARBRUL analysis, as in any scientific endeavor, is to develop the most parsimonious model that still accounts for the data. To achieve this goal, we need to test whether the results are statistically significant or whether there is a good likelihood that they might be due to chance. In VARBRUL analysis, achieving the most parsimonious model involves testing whether entire factor groups significantly contribute to the overall goodness-of-fit of the model and testing whether factors within groups differ significantly from one another. Naturally, factors should only be combined where there is linguistic or social justification for doing so. Guy (1980), for example, in a study of -t,d deletion, found that the VARBRUL weights for regular past tense verbs and past participles did not differ significantly from one another. As discussed in more detail below, he combined these forms on the basis of their common underlying morphological structure. It would have made little sense, however, to combine following consonants and vowels, for example, which not only contrast phonologically but have been found in other studies to differ significantly in their effect on deletion.

3.1 Significance testing with VARBRUL

VARBRUL provides a means of testing whether a particular factor group contributes significantly to the model of variation by means of *step-up/step-down analysis*. This involves performing a run with only one factor group and then adding each of the other factor groups to the analysis, one at a time, until all factor groups are included. When the full model with all factor groups is reached, VARBRUL then removes one factor group at a time until only one remains. During each individual run, the factor weights and the log-likelihood are calculated. At the end of the analysis, the program outputs a file with the details of each run and an indication of the best stepping-up run and the best stepping-down run. The factor groups included in the best stepping-up and

stepping-down runs should be the same. These are the factor groups that are significant at or above the $p < .05$ level. Factor groups that are not included in these runs do not contribute significantly to the variation. The factor weights calculated during these runs are used to report the results of the study.

In addition to testing the significance of factor groups, it is also necessary to test whether individual factors within groups differ significantly from one another. This calculation is done by comparing the log likelihoods of two VARBRUL runs, one with the factor coded as is, and one with a recode in which the factor is eliminated or collapsed with another factor. The following test is used in order to determine whether the difference between the VARBRUL weights is significant:

$$\chi^2_{\nu\alpha} \approx -2 \text{ (log likelihood}_1 - \text{log likelihood}_2)$$

That is, twice the difference in the log likelihoods of the separate analyses performed (1) with and (2) without the factor in question asymptotically approximates a chi-square distribution where ν is the number of degrees of freedom and α is the probability that the effect attributed to the factor in question is greater than would be expected by chance. The degrees of freedom used in calculating the above chi-square statistic are the difference between the degrees of freedom in the two runs. If only one factor is eliminated, there will only be a single degree of freedom used in testing for the significance of that factor.

3.2 Interpreting the results of VARBRUL analysis

VARBRUL enables us to give precise and replicable measures of the strength of a wide range of contextual influences on the choice among variable linguistic forms. However, simply reporting results is not sufficient. Rather, our goal is to understand why we achieve the results that we do. Take the effect of grammatical category on the likelihood of -t,d deletion as an example.

A number of explanations have been proposed for the pattern for grammatical function seen in Wolfram (1969), Guy (1980), Baugh (1983), and numerous other studies. At first glance, it appears that the functional load carried by final -t,d might provide an adequate explanation. Nothing is lost if a speaker says *jus' me* instead of *just me*, but it is not so easy to determine whether *I miss/Ø/ my friend* refers to a missed past appointment or to an ongoing emotional state. As we have seen, however, Guy (1980), showed that the rate of -t,d deletion from past participles, e.g. *she was miss/Ø/ by all*, did not differ significantly from the rate of deletion from past tense forms, despite the fact that past participles carry a lighter functional load.

Guy's (1980) finding that -t,d is deleted from past participles and past tense verbs at the same rate suggests that we must look beyond functionalism for an explanation of the ordering of grammatical constraints. A number of possible

explanations have been proposed. Guy (1993), for example, observed that the grammatical categories that are subject to -t,d deletion are characterized by different internal morphological boundaries, and that regular past tense forms and past participles have the same internal structure. The results for grammatical category can thus be explained by a boundary constraint on -t,d deletion. A deletion rule applies freely when no internal boundary is present, as is the case with monomorphemes such as *past*. Deletion is inhibited somewhat by the formative boundary in semi-weak verbs, and strongly inhibited by the inflectional boundary in regular past tense verbs and past participles.

Other explanations have also been advanced (see Labov 1997 for a full account). Guy (1991) proposed an exponential model of constraints to explain the relationships observed in the grammatical category factor group, which related the retention of past tense, semi-weak, and monomorphemic clusters in the ratio of x: x^2: x^3. He explained this ratio as a consequence of the multilevel architecture of lexical phonology (Kiparsky 1985), whereby the three types of clusters are subject to one, two, or three passes of a deletion rule. Two subsequent studies (Bayley 1997, Santa Ana 1992) confirmed the predictions of the exponential model. More recently, Kiparsky (1994) suggested that the exponential relationship pointed out by Guy could be explained by an exploded optimality constraint.

The purpose here is not to argue which of these explanations is correct. The point is to demonstrate that the results achieved by the use of VARBRUL – or any other statistical program – do not in and of themselves provide explanations about linguistic structure or the meaning of the social distribution of linguistic variants. Rather, explanations must be sought in linguistic theory and in our understanding of the history and social structure of the communities we study.

3.3 Limitations of VARBRUL

VARBRUL has proven to be an extremely productive tool for the study of linguistic variation. However, it does suffer from a number of limitations. First, GoldVarb, the Macintosh version of the program dates from 1991, and a multinomial version has not been implemented for the Macintosh operating system. Also, until very recently, PC users had to resort to Susan Pintzuk's DOS version dating from 1988. Recently, however, Sali Tagliamonte, Pintzuk, and other researchers at the University of York have begun to develop a Windows version of the program, scheduled for release in 2001.[5] Second, because the use of the program has generally been restricted to research in sociolinguistics and second language acquisition (by researchers who have also been trained in sociolinguistics), VARBRUL users often find themselves without support. As Young and Yandell (1999) note, statistical consultants at most universities are not aware of VARBRUL. Third, as noted in the previous section, VARBRUL does not provide a convenient way to test for interactions among factor groups. As Sankoff (1988) observed, in the case of properly defined

linguistic factors, interaction is not normally a problem. To return to the example of -t,d deletion in English, clearly there is no a priori reason to suspect that grammatical function and the following segment, or syllable stress and the preceding segment, will interact. However, as noted above, non-linguistic factors such as social class, gender, and ethnicity often interact. It is possible – and sometimes desirable – to control such interactions within VARBRUL by recoding. Alternately, one might consider each participant in the study as a separate factor and, assuming that the variable under investigation has social significance, use the VARBRUL weights to group individuals into relevant social categories.[6] Whether such procedures are desirable, however, depends on the goals of the study. If potential interactions among various social factors are a major interest of the investigation, VARBRUL is not the most suitable tool.

4 General Logistic Regression Models

In recent years, a number of researchers have used the logistic regression modules in commercially available statistical packages to overcome some of the limitations of VARBRUL.[7] Berdan (1996), for example, used the logistic regression module in SPSS to reanalyze Schumann's (1978) longitudinal data on the acquisition of English negation by "Alberto," a Costa Rican immigrant to the United States. Schumann's original study, which has been influential in second language acquisition research, provided the basis for his acculturation model and for the concept of fossilization. According to Schumann, Alberto showed very little progress in the acquisition of negation. He concluded that during the course of the study, which lasted a year, his subject "remained in the first [no + V] stage" (1978: 65).

In his reanalysis, Berdan was able to show that Alberto did in fact make progress toward the target language. Although he used *no + V* to express negation throughout the year, this most basic structure alternated with the more target-like unanalyzed *dont*, which became more frequent over time. Berdan used a general logistic regression model rather than VARBRUL for two main reasons. First, the general model enables the researcher to represent independent variables as continuous, "[a] procedure [that] allows for an intuitive representation of time – the variable that is integral to the modeling of learning or language acquisition" (Berdan 1996: 212). Second, the logistic regression model allows the researcher to calculate the main effects for independent variables (factors in VARBRUL) as well as to calculate interactions among them. Since the interaction of linguistic variables with time and with style was a main focus of Berdan's study, the general logistic model proved the more useful tool.

Young and Yandell (1999) provide an example of a direct comparison of results from a general logistic regression model and from VARBRUL. The comparison was undertaken in response to Saito's (1999) critique of Young's

(1991) original VARBRUL analysis of -s plural variation in the interlanguage of Chinese learners of English. Saito criticized Young's study for grouping speakers together by proficiency level rather than considering them individually and for failing to consider possible interactions between proficiency and the preceding segment effect. The first criticism is not germane to the choice between general logistic regression models and VARBRUL because it is a simple matter to include individuals as factors, regardless of the statistical package used. The second criticism is perhaps more relevant because, as we have seen, VARBRUL does not provide an easy way to deal with interactions. In the original study, Young (1991) found that proficiency level was statistically significant when all of the data were run together, with high proficiency learners more likely to mark -s plurals than low proficiency learners. He subsequently ran two separate analyses by proficiency level and found that the preceding segment, animacy, and definiteness affected high and low proficiency learners differently.

Young and Yandell used the GENMOD procedure in the SAS/STAT software package (SAS Institute 1996) to reanalyze the original data. As shown in table 5.4, the results are comparable to the original analysis. Both the original VARBRUL analysis and the GENMOD analysis showed that proficiency, noun position, syntactic function, preceding and following segments, and redundancy significantly affected speakers' use of plural -s. In addition, the GENMOD analysis showed significant interactions between proficiency and definiteness, animacy, and the preceding segment.

As we see in table 5.4, for example, both VARBRUL and GENMOD show that redundant plural marking in the NP favors -s plural marking for speakers at all levels of English proficiency, as do prenominal modifiers. When all speakers are included in the analysis, following vowels favor -s plural marking. However, this is a rather weak constraint, as indicated by the fact that it failed to reach statistical significance when separate analyses were run by proficiency level. Table 5.4 also allows us to compare the interactions found in the GENMOD analysis with the results of the separate VARBRUL analyses by proficiency level. For example, in the GENMOD analysis, proficiency by animacy is significant because, as revealed in the VARBRUL analyses by proficiency level, the animacy of the NP affects high and low proficiency speakers differently. For high proficiency speakers, animate NPs favor -s plural marking. For low proficiency speakers, animate NPs disfavor plural marking.

5 Multivariate Analysis: Summary and Conclusions

Numerous examples from studies of linguistic variation conducted around the world have shown that multivariate analysis is necessary if we are to understand the complex array of factors that may potentially influence the choice of

Table 5.4 Comparison of VARBRUL and GENMOD results for the analysis of -s plural marking in Chinese-English interlanguage

Factor	VARBRUL: all speakers	VARBRUL: low proficiency	VARBRUL: high proficiency	GENMOD
Proficiency	High proficiency favors	na	na	na
Individual speaker	na	na	na	significant
Redundant plural marking in NP	Redundant marking favors	Redundant marking favors	Redundant marking favors	Redundant marking favors
Syntactic function of NP	Adverbial favors	Adverbial favors	Adverbial favors	Adverbial favors
Position of noun within NP	Prenominal modifier favors	Prenominal modifier favors	Prenominal modifier favors	Prenominal modifier favors
Preceding segment	Vowels, stops, and non-sibilant fricatives favor	Vowels and stops favor	ns	Factor order is fricative > stop > vowel > nasal > sibilant > lateral
Following segment	Vowels favor	ns	*ns*	Vowels favor
Animacy	ns	disfavors	favors	ns
Definiteness	ns	ns	disfavors	ns
Proficiency x definiteness	na	na	na	significant
Proficiency x animacy	na	na	na	significant
Proficiency x preceding segment	na	na	na	significant

na = not applicable, either because the factor group was not included in the model (e.g. the interaction factor groups in the VARBRUL analyses) or because the factors in the group were dealt with elsewhere in the model (e.g. when speakers were coded as individual factors, the proficiency factor group was no longer relevant). ns = not significant at $p < 0.05$.
Sources: Young (1991: 144–5), Young and Yandell (1999)

one or another linguistic variant and the systematicity that often underlies variable language production. The choice then, is not whether to do multivariate analysis, but which particular model to use. Within the quantitative tradition, VARBRUL has generally been the preferred tool. As Berdan notes, "VARBRUL has . . . proven to be a powerful analytic device for identifying significant linguistic, social, and interactional factors that differentiate or condition probabilities associated with linguistic variables" (1996: 209). Nevertheless, although VARBRUL can serve effectively to model variation among linguistic factors, it is not appropriate where interaction, either among non-linguistic factors, or between non-linguistic and linguistic factors, is a main focus of the study. In that case, a more general logistic regression model is preferable. Finally, there is the question of audience. Most researchers trained in quantitative sociolinguistics are familiar with VARBRUL, at least to the extent of being able to interpret the results presented in variationist studies. However, the program and the terminology associated with it are unfamiliar to readers outside the field of quantitative sociolinguistics. For this reason, Young and Yandell (1999) recommend the use of widely available programs such as the GENMOD procedure in SAS/STAT for the analysis of interlanguage variation, where the audience is drawn from many disciplines.

6 Future Directions: Variationist Ethnography

Books by Chambers (1995) and Wolfram and Schilling-Estes (1998) review in detail many of the achievements of more than three and a half decades of research within the quantitative paradigm. It is now beyond dispute that much of the variation in language that was previously thought to be random is indeed systematic, and that eloquence, logic, and clarity of expression are not the particular properties of standard languages. Moreover, although public attitudes have been slow to change, work on socially stigmatized varieties, particularly the varieties used by ethnic minorities and members of the working class, has served as important evidence to combat popular misperceptions of such varieties as being illogical and their speakers as incapable of mastering national or regional standard varieties. On the contrary, sociolinguistic analysis has revealed beyond any doubt that these varieties are orderly, complex, and complete linguistic systems. In this respect, Labov's (1969b) seminal essay, "The logic of nonstandard English," has proven particularly important. The work of Geneva Smitherman, Labov, and other scholars in establishing the linguistic rights of African-American school children (Labov 1982), at least in one area of the United States, provides an example of the influence of sociolinguistics in the judicial arena. More recently, Baugh (2000), Rickford (1999), and many others in the field have sought to resolve some of the public confusion and to combat the racist stereotypes surrounding the Oakland, California school board's attempt to use AAVE to teach standard English. Finally, it is well-established

that synchronic variation provides a key to understanding language change (Milroy 1992).

Despite these substantial achievements, work in the quantitative paradigm has focused less on exploring the meaning of linguistic variation to the members of the communities studied than on demonstrating the systematic nature of variability at an abstract level and establishing correlations among linguistic variables and traditional social categories such as age and class. Studies are beginning to appear, however, that not only demonstrate the correlations between the use of linguistic variables and social categories, but also show how speakers deploy their linguistic resources, along with other symbolic resources, to construct and reinforce the social categories to which they belong. Eckert's (1989, 2000) work on Euro-American high school students in the Detroit area and Mendoza-Denton's (1997) study of Chicana and Mexican immigrant adolescents in California are two important examples of this direction in research within the quantitative paradigm.

Traditionally, variationist studies have grouped participants according to pre-determined social categories such as class, ethnicity, gender, and age,[8] and examined possible correlations between these non-linguistic factors and use of socially salient linguistic variables. Rather than grouping participants by pre-determined social categories, Eckert and Mendoza-Denton, through intensive ethnographic investigation, sought to discover the social categories that participants themselves found meaningful. At times, these categories overlapped with the categories usually considered in variationist studies, but at other times they differed considerably. For example, Mendoza-Denton found six distinct groups among the Mexican immigrant and Chicana students in a California high school, ranging from immigrant *piporras* ("country girls"), who tended to preserve traditional rural Mexican values, to the mostly non-immigrant "Latina jocks," who participated in the school culture, especially sports, and tended to accept the values of the larger society that the school represented. Moreover, different group affiliation was associated with differing patterns of language use, including different VARBRUL weights for raising and lowering of /ɪ/, the variable Mendoza-Denton examined in detail. Clearly, much would have been lost by simply grouping speakers according to ethnicity or immigrant generation.

Although ethnographically oriented studies of variation such as Eckert (2000) and Mendoza-Denton (1997) have been relatively unusual, they are certainly not unprecedented. As Eckert points out, the concern with local identity and participants' views of that identity was a central focus of one of the earliest studies in the quantitative tradition, Labov's (1963) examination of centralization of (ay) and (aw) by residents of Martha's Vineyard in Massachusetts. In this sense, then, some of the more interesting recent work within the quantitative paradigm represents a return to the roots of the discipline.

In summary, quantitative analysis has enabled us to obtain numerous insights into linguistic structure, the social meaning of linguistic variation, and the nature of language change. From a social perspective, the methods of multivariate

analysis developed in sociolinguistics have been particularly important in demonstrating the systematic nature of stigmatized language varieties, including AAVE, Montreal French, popular Puerto Rican and US Spanish, as well as many others. Current work, in which quantitative analysis is informed by ethnographic fieldwork, promises further insights into the ways in which language users employ variation to construct social identities. Finally, thanks to the widespread availability of powerful statistical software packages, sociolinguists now have many options at their disposal. As the field becomes more experienced in quantitative methods, and particularly in the range of available multivariate applications, new creative possibilities for quantitative analysis will doubtless open up.

ACKNOWLEDGMENTS

Work on this chapter was supported by a University of Texas at San Antonio Faculty Research Award. Many of the ideas expressed here have been developed over the years through conversations with Gregory Guy and Dennis Preston, and through collaboration with Richard Young, with whom I co-authored a guide to VARBRUL analysis (Young and Bayley, 1996) and conducted a number of workshops on quantitative methods. The chapter also benefited from Jack Chambers' careful editing and numerous helpful suggestions. Faults that remain, of course, are my own.

NOTES

1 A number of researchers, including Berdan (1996), Fasold (1991), Guy (1993), and Sankoff (1988), have pointed out that much of the work done as "variable rule analysis" does not involve rules of the type illustrated here. As Sankoff (1988) observed, however, the statistical analysis does not depend on the origin of the variation in the data.

2 GoldVarb 2.1 and a brief manual may be downloaded from David Sankoff's web page: www.crm.umontreal.ca/~sankoff/GoldVarb_Eng.html. VARBRUL for MS-DOS is available by anonymous ftp at the University of Pennsylvania by setting your web browser to: ftp://ftp.cis.upenn.edu/pub/ldc/misc_sw/varbrul.tar.Z.

3 The issue of whether copula absence represents a deletion or an insertion rule is important to debates about the origin of AAVE. Evidence that copula absence is the result of an insertion rule would support a creolist position on the origin of AAVE. On the other hand, evidence for a deletion rule, whereby copula deletion follows copula contraction, as Labov (1969a) proposed, would support the position that AAVE is essentially similar to other dialects of English.

4 Guy (1988) discusses the issue of dealing with interactions in detail. See also Bayley et al. (2000) for one solution to the problem of interaction.

5 For up-to-date information on VARBRUL for Windows, contact Dr. Sali Tagliamonte at sali.tagliamonte@utoronto.ca.
6 When more than one linguistic variable is involved, and the researcher wishes to investigate the relationship between individual use of a number of potentially related linguistic variables (e.g. -t,d deletion, alveolization of /ŋ/, and negative concord) and social distinctions, principal components analysis is an effective method. See Horvath and Sankoff's (1987) study of the Sydney speech community for an example of the use of this method.
7 See Reitveld and van Hout (1993) for information on the use and interpretation of general logistic regression models.
8 There are, of course, important exceptions to this traditional practice, including Labov's (1972a) work with the Jets and the Cobras in New York. In addition, Gillian Sankoff (1980, 1989) has long combined ethnographic and variationist approaches in her work in Montreal and New Guinea. The social network approach developed by James and Lesley Milroy represents another exception.

REFERENCES

Baugh, John (1983). *Black Street Speech: Its History, Structure, and Survival*. Austin, TX: University of Texas Press.

Baugh, John (2000). *Beyond Ebonics: Linguistic Pride and Racial Prejudice*. New York: Oxford University Press.

Bayley, Robert (1994a). Consonant cluster reduction in Tejano English. *Language Variation and Change* 6: 303–26.

Bayley, Robert (1994b). Interlanguage variation and the quantitative paradigm. In Elaine Tarone, Susan M. Gass and Andrew Cohen (eds.), *Research Methodology in Second-language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum. 157–81.

Bayley, Robert (1997). Variation in Tejano English: Evidence for variable lexical phonology. In Cynthia Bernstein, Thomas Nunnally and Robin Sabino (eds.), *Language Variety in the South Revisited*. Tuscaloosa: University of Alabama Press. 197–209.

Bayley, Robert, Ceil Lucas and Mary Rose (2000). Variation in American Sign Language: The case of DEAF. *Journal of Sociolinguistics* 4: 81–107.

Bayley, Robert and Lucinda Pease-Alvarez (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9: 349–71.

Beebe, Leslie M. (1977). The influence of the listener on code-switching. *Language Learning* 27: 331–9.

Berdan, Robert (1996). Disentangling language acquisition from language variation. In Robert Bayley and Dennis R. Preston (eds.), *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins. 203–44.

Cedergren, Henrietta (1973). The interplay of social and linguistic

factors in Panama. Ph.D. dissertation, Cornell University.
Cedergren, Henrietta and David Sankoff (1974). Variable rules: Performance as a statistical reflection of competence. *Language* 50: 233–55.
Chambers, J. K. (1995). *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Blackwell.
Eckert, Penelope (1989). *Jocks and Burnouts: Social Categories and Identity in the High School*. New York: Teachers' College Press.
Eckert, Penelope (2000). *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. Oxford: Blackwell.
Fasold, Ralph (1991). The quiet demise of variable rules. *American Speech* 66: 3–21.
Feagin, Crawford (1979). *Variation and Change in Alabama English: A Sociolinguistic Study of the White Community*. Washington, DC: Georgetown University Press.
Guy, Gregory R. (1980). Variation in the group and in the individual: The case of final stop deletion. In William Labov (ed.), *Locating Language in Time and Space*. New York: Academic. 1–36.
Guy, Gregory R. (1981). Linguistic variation in Brazilian Portuguese: Aspects of the phonology, syntax, and language history. Ph.D. dissertation, University of Pennsylvania.
Guy, Gregory R. (1988). Advanced VARBRUL analysis. In Kathleen Ferrara, Becky Brown, Keith Walters and John Baugh. (eds.), *Linguistic Contact and Change: Proceedings of the Sixteenth Annual Conference on New Ways of Analyzing Variation*. Austin, TX: University of Texas Department of Linguistics. 124–36.
Guy, Gregory R. (1991). Explanation in variable phonology. *Language Variation and Change* 3: 1–22.
Guy, Gregory R. (1993). Quantitative analysis. In Dennis R. Preston (ed.), *American Dialect Research*. Amsterdam: John Benjamins. 223–49.
Guy, Gregory R. (1997). The grammar of variation. In Frans Hinskens, Roeland van Hout and W. Leo Wetzels (eds.), *Variation, Change and Phonological Theory*. Amsterdam: John Benjamins. 125–43.
Horvath, Barbara and David Sankoff (1987). Delimiting the Sydney speech community. *Language in Society* 16: 179–204.
Kiparsky, Paul (1985). Some consequences of lexical phonology. In Colin McEwen and James Anderson (eds.), *Phonology Yearbook 2*. Cambridge: Cambridge University Press. 85–138.
Kiparsky, Paul (1994). An optimality-theoretic perspective on variable rules. Paper presented at NWAVE 23, Stanford University, California.
Labov, William (1963). The social motivation of a sound change. *Word* 19: 273–307.
Labov, William (1966). *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
Labov, William (1969a). Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715–62.
Labov, William (1969b). The logic of nonstandard English. *Georgetown Monographs in Languages and Linguistics* 22: 1–44.
Labov, William (1972a). *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
Labov, William (1972b). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William (1978) Where does the sociolinguistic variable stop? A reply to B. Lavendera. *Texas Working Papers in Sociolinguistics* 44. Austin, TX: SW Educational Development Laboratory.

Labov, William (1982). Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor. *Language in Society* 11: 165–201.

Labov, William (1989). The child as linguistic historian. *Language Variation and Change* 1: 85–97.

Labov, William (1994). *Principles of Linguistic Change*, vol. 1: *Internal factors*. Oxford: Blackwell.

Labov, William (1997). Resyllabification. In Frans Hinskens, Roeland van Hout and W. Leo Wetzels (eds.), *Variation, Change and Phonological Theory*. Amsterdam: John Benjamins. 145–79.

Labov, William, Paul Cohen, Clarence Robins and John Lewis, (1968). *A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City*. Philadelphia: US Regional Survey.

Lavendera, Beatrice (1978). Where does the sociolinguistic variable stop? *Language in Society* 7: 171–82.

Liddell, Scott K. and Robert E. Johnson (1989). American Sign Language: The phonological base. *Sign Language Studies* 64: 195–278.

Lucas, Ceil (1995). Sociolinguistic variation in ASL: The case of DEAF. In Ceil Lucas (ed.), *Sociolinguistics in Deaf Communities*, vol. 1. Washington, DC: Gallaudet University Press. 3–25.

Mendoza-Denton, Norma (1997). Chicana/Mexican identity and linguistic variation: An ethnographic and sociolinguistic study of gang affiliation in an urban high school. Ph.D. dissertation, Stanford University, California.

Milroy, James (1992). *Linguistic Variation and Change*. Oxford: Blackwell.

Milroy, Lesley (1987). *Language and Social Networks*, 2nd edn. Oxford: Blackwell.

Norušis, Marija J. and SPSS Inc. (1996). *SPSS Advanced Statistics 6.1*. Chicago: SPSS Inc.

Pintzuk, Susan (1988). VARBRUL programs for MS-DOS [computer software]. Philadelphia: University of Pennsylvania Department of Linguistics.

Poplack, Shana (1980). "Sometimes I'll begin a sentence in Spanish and *termino in español*": Towards a typology of code-switching. *Linguistics* 18: 581–618.

Poplack, Shana and Svitlana Budzhak-Jones (1998). The visible loanword: Processes of integration seen in bare English-origin nouns in Ukrainian. In Vera Regan (ed.), *Contemporary Approaches to Second Language Acquisition in Social Context*. Dublin: University College Dublin Press. 137–67.

Rand, David and David Sankoff (1991). *GoldVarb: A Variable Rule Application for the Macintosh* (version 2.1) [computer program]. Montreal: Centre de recherches mathématiques, Université de Montréal.

Reitveld, Toni and Roeland van Hout (1993). *Statistical Techniques in the Study of Language and Language Behavior*. Berlin: de Gruyter.

Rickford, John R. (1987). *Dimensions of a Creole Continuum: History, Texts, and Linguistic Analysis of Guyanese Creole*. Stanford, CA: Stanford University Press.

Rickford, John R. (1999). The Ebonics controversy in my backyard: A sociolinguist's experiences and reflections. *Journal of Sociolinguistics* 3: 267–75.

Rickford, John R. Arnetha Ball, Renée Blake, Raina Jackson and Naomi Martin (1991). Rappin on the copula coffin: Theoretical and methodological issues in the analysis of copula variation in African-American Vernacular English. *Language Variation and Change* 3: 103–32.

Roberts, Julie (1997). Acquisition of variable rules: A study of (-t,d) deletion in preschool children. *Journal of Child Language* 24: 351–72.

Rousseau, Pascale (1989). A versatile program for the analysis of sociolinguistic data. In Ralph W. Fasold and Deborah Schiffrin (eds.), *Language Change and Variation*. Amsterdam: John Benjamins. 395–409.

Rousseau, Pascale and David Sankoff (1978). Advances in variable rule methodology. In David Sankoff (ed.), *Linguistic Variation: Models and Methods*. New York: Academic. 57–69.

Saito, Hidetoshi (1999). Dependence and interaction in frequency data analysis in SLA research. *Studies in Second Language Acquisition* 21: 453–75.

Sankoff, David (1988). Variable rules. In Ulrich Ammon, Norbert Dittmar and Klaus J. Mattheier (eds.), *Sociolinguistics: An International Handbook of the Science of Language and Society*, vol. 2. Berlin: de Gruyter. 984–97.

Sankoff, David and William Labov (1979). On the uses of variable rules. *Language in Society* 9: 189–222.

Sankoff, Gillian (1980). *The Social Life of Language*. Philadelphia: University of Pennsylvania Press.

Sankoff, Gillian (1989). A quantitative paradigm for the study of communicative competence. In Richard Bauman and Joel Sherzer (eds.), *Explorations in the Ethnography of Speaking*, 2nd edn. Cambridge: Cambridge University Press. 18–49.

Sankoff, Gillian (1990). The grammaticalization of tense and aspect in Tok Pisin and Sranan. *Language Variation and Change* 2: 295–312.

Santa Ana A., Otto (1992). Chicano English evidence for the exponential hypothesis: A variable rule pervades lexical phonology. *Language Variation and Change* 4: 275–89.

SAS Institute Inc. (1996). *SAS/STAT Software: Changes and Enhancements through Release 6.11*. Cary, NC: SAS Institute Inc.

Schiffrin, Deborah (1982). Tense variation in narrative. *Language* 57: 45–62.

Schumann, John H. (1978). *The Pidginization Process: A model for Second Language Acquisition*. Rowley, MA: Newbury House.

Tarone, Elaine (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning* 35: 373–404.

Trudgill, Peter (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Weiner, E. Judith and William Labov (1983). Constraints on the agentless passive. *Journal of Linguistics* 19: 29–58.

Weinreich, Uriel, William Labov and Marvin Herzog (1968). Empirical foundations for a theory of language change. In Winfred P. Lehmann and Yakov Malkiel (eds.), *Directions for Historical Linguistics: A Symposium*. Austin, TX: University of Texas Press. 95–188.

Wolfram, Walt (1969). *A Sociolinguistic Description of Detroit Negro Speech*. Washington, DC: Center for Applied Linguistics.

Wolfram, Walt and Ralph W. Fasold (1974). *The Study of Social Dialects in*

the United States. Englewood Cliffs, NJ: Prentice-Hall.
Wolfram, Walt and Natalie Schilling-Estes (1998). *American English: Dialects and Variation*. Oxford: Blackwell.
Young, Richard (1991). *Variation in Interlanguage Morphology*. New York: Peter Lang.
Young, Richard and Robert Bayley (1996). VARBRUL analysis for second language acquisition research. In Robert Bayley and Dennis R. Preston (eds.), *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins. 253–306.
Young, Richard and Brian Yandell (1999). Top-down versus bottom-up analyses of interlanguage data: A reply to Saito. *Studies in Second Language Acquisition* 21: 477–88.