

Variation and Language, an Overview

in

Encyclopedia of Language and Linguistics 2nd Edition

Walt Wolfram
North Carolina State University
English Department, Box 8105
Raleigh, NC 27695-8105
USA

ABSTRACT

The past half-century has witnessed remarkable growth in the study of language variation, and it has now become a highly productive subfield of research in sociolinguistics. This section considers the locus of variability within language and discusses the linguistic variable as a heuristic construct in the analysis and the description of language variation. It further discusses the nature of systematic language variation, as well the role of structured heterogeneity in language change. Finally, it considers possible claims about representing language capability with respect to language variation and implications for a grammar of language.

THE LOCUS OF VARIATION

If structure is at the heart of language, then variation defines its soul. As Sapir (1921: 147) put it, “Everyone knows that language is variable.” Furthermore, variation allows us to differentiate individuals, groups, communities, states and nations. Notwithstanding the pervasive nature of variability in language, it has often been disregarded or dismissed as tangential to the description of structural patterning and irrelevant to the study of linguistic competence. In fact, it was not until the advent of sociolinguistics a half-century ago that the admission of language variation became more than a footnote to linguistic description. The study of language variation is now one of the most rapidly expanding subfields of linguistics with a well-established cohort of researchers, regular conferences, and scholarly journals, but its status is still somewhat marginal within theoretical linguistics, notwithstanding the insistence of William Labov that the study of language variation is central to the solution of fundamental problems in linguistic theory (e.g. Labov 1966, 2001).

Variability is everywhere in language, from the unique details in each production of a sound or sign to the auditory or visual processing of the linguistic signal. In fact, one of the amazing facts about human communication is the demonstrated ability to normalize the inherent variation within every spoken or signed message in processing the linguistic signal. Though language variation is persistent and pervasive, it is not all equally interesting, even to those who focus on the systematic nature of language variation, so-called LANGUAGE VARIATIONISTS. For example, there is considerable variation in speech or sign production related to the physical make-up of an individual speaker or signer—differences in vocal tract size in spoken language or differences in the size of the hands and body used in signing—that is of interest to those who study language normalization, but this is not generally the focus of systematic language variation studies. Similarly, the relative fluency of production related to idiosyncratic behavior is not of

concern to those interested in language variation, though the dichotomy between meaningful social differences and socially insignificant personal differences is not always clear-cut. Generally speaking, interest in language variation focuses on differences that have some social significance in terms of group behavior rather than personal idiosyncrasies, though socially meaningful aspects of individual speaker performance are of interest to those interested in language variation.

The empirical reality is that the boundaries of significant and insignificant language variation are often gradient and obscure rather than discrete and transparent. There is, for example, a fine line between consequential and inconsequential fluctuation in the vowel formants of repeated productions of a given vowel in a constant phonetic context. Some variability in each vowel production is expected, but a larger envelope of fluctuation in the formants might be indicative of a significant vowel change in progress. Determining significance for an envelope of vowel formants, however, can only be arrived at through a series of exploratory procedures that takes into account individual speaker traits, social group norms, the production of other vowels in the system, and so forth. Obviously, the empirical boundaries of investigation for language variation studies are not as straightforward as they may appear in descriptions of systematic variability found in the research literature and often involve preliminary decisions about what to include and what to exclude in the study of orderly variation.

In traditional linguistic description the notion of variation within structural units has often been acknowledged under labels such as “free fluctuation,” “optional rules,” and “free variants.” Though long recognized, this variability has nonetheless been considered to be “an area of little importance” (Crystal 2003: 189), a kind of garbage heap for variants that could not be predicted invariably within the categorical framework assumed under most models of linguistic description. But in keeping with the adage that “one person’s garbage is another person’s treasure,” the examination of such variation has become the cornerstone of sociolinguistics, with several professional journals dealing regularly with issues of language variation, including one dedicated exclusively to this issue (*Language Variation and Change*), and prominent sociolinguistics conferences that routinely feature presentations on language variation. This includes perhaps the most influential sociolinguistics conference in the world, the annual conference New Ways of Analyzing Variation (NWAV), now well into its third decade of continuous growth.

THE LINGUISTIC VARIABLE

An essential construct in the study of linguistic variation is the LINGUISTIC VARIABLE, a structural unit that includes a set of fluctuating variants showing meaningful co-variation with an independent set of variables. The linguistic variable was explicitly set forth in early variation studies by William Labov (1966), the acknowledged founder of the field of language variation studies, though this construct was certainly implicit in work previous to that point. On one level, the relationship of the variable to its variants may be likened to that of the classic emic-etic relationship in linguistic description, such as that between a morpheme and its allomorphs or a phoneme and its allophones; however, as we shall see, it is hardly confined by such structural relationships and boundaries. The unit is an abstraction since linguistic variables have no concrete reality apart from the particular variants through which they are realized.

Operationally, the linguistic variable has been used to encompass a wide range of fluctuating variants. The set may be a structural type, such as grammatical category (e.g. a tense category, possessive category), a phoneme (e.g. the phoneme /θ/ or the phoneme /ŋ/ in English),

or a natural class of items in a specific linguistic context (e.g. coronal stops in syllable-coda position). It may also be defined in terms of the application of a general process, such as the use of a particular type of contraction (e.g. English variation between *She has not* vs. *She hasn't* vs. *She's not*) or in terms of a syntactic relationship, such as concord (e.g. negative concord in English as in *He didn't do anything about any problem* vs. *He didn't do nothing about no problem*) or phrasal constituency (e.g. head-initial vs. head-final phrases). Variants are usually established apart from theoretical models of language description and may, for all intents and purposes, be accommodated within any syntactic model (e.g. Principles and Parameters, Minimalism, Head-Driven Phrase Structure Grammar, Lexical Functional Grammar, and Construction Grammar) or phonological model (e.g. Generative Phonology, Lexical Phonology, Optimality Theory, Exemplar Theory), though there are significant implications for how such variation may be accommodated within a particular model of language. Finally, the linguistic variable may be defined as a simple lexical choice (e.g. *soda*, *pop*, *tonic*, or *co(ca)-cola* for a carbonated drink in various regions of the US) or even in terms of a speech act (e.g. alternatives for expressing a 'directive' or a 'compliment').

As noted, the linguistic variable can cover a full range of linguistic units and relationships. By the same token, the relationship of the variants to each other is not necessarily limited by the fundamental typology operative in linguistic analysis and description. For example, a given phonological variable may encompass allophonic fluctuation within a given phoneme, different phonemes, or combine allophonic and phonemically distinct variants within a unitary variable. Thus, the pronunciation of the vowel of English words like *coffee* or *caught* may involve variants that range from different phonetic productions within the same phonemic vowel (e.g. [kaofi] vs. [koəfi]) to a phonemic merger with another phonemic unit (e.g. the vowel of *caught* and *cot* are both pronounced as [ɒ]). In this case, the variants appear to be united by the fact that they exist within a single lexical item. Furthermore, in some applications of the linguistic variable, variants may involve competing systems or grammars (Kroch 1989) so that variation cannot even be considered inherent within a unitary system. Under this interpretation, typological correspondence or equivalency in the competing systems is the apparent basis for including different variants in a linguistic variable.

The notion of the linguistic variable has been applied to a full range of levels within language, though not without some argument about its appropriate application. Phonological and morphosyntactic variation have tended to dominate language variation research in synchronic and apparent time studies while syntactic variation has been a significant locus of investigation in diachronic studies. For synchronic and apparent time studies, which have tended to use natural conversational interviews as their primary databases, there is both a practical and theoretical explanation for the focus on phonology and morphosyntax. In natural conversation, syntactic phenomena simply do not occur at sufficient frequency levels for meaningful quantitative analyses of systematic variability, the procedural underpinning for variation studies. But there have also been theoretical concerns about the application of variation studies beyond phonology related to the assumption of "semantic equivalency", that is, the notion that there is no change in referential meaning based on the selection of one variant or another (Lavandera 1978). While it may safely be assumed that the variant pronunciations of unstressed *-ing* in the English word *running* as a final velar [ŋ] or coronal nasal [n] have the same denotational referent, questions about semantic equivalency have arisen in examining variation on other levels. For example, it may be argued that there are subtle meaning differences in the Montreal French use of variants *être* and *avoir* in different contexts (Sankoff & Thibault 1977). Similarly, it may be questioned

whether the difference between the ‘active’, ‘*be* passive’, and ‘*get* passive’ in English (e.g. *They broke into the store* vs. *The store was broken into* vs. *The store got broken into*) are authentic alternatives for “saying the same thing” (Weiner & Labov 1983). Lavandera (1978) has suggested that the condition of strict semantic equivalency need not be applied rigorously so that the linguistic variable could be used in study language variation beyond phonology, a guideline now generally adopted in studies of syntactic variables.

The focus on syntax in diachronic studies of language variation is, to a large extent, a product of the kinds of available texts for investigation. Historical documents tend to be more reliable for examining grammatical than phonetic and phonological variation. Furthermore, the increasing accessibility of mega-corpora with automated search capabilities for historical texts has made the investigation of syntactic variability much more accessible to variation researchers.

Though there are lingering questions about the precise structural status of the linguistic variable, it remains a primitive construct in the study of language variation. Most of the lingering questions and qualifications have to do with its precise status as a structural unit and its possible relationship to a grammar of language rather than its usefulness as convenient heuristic in the investigation of language variation. Whether or not it is explicitly recognized, the linguistic variable remains at the procedural foundation of variation analysis, since one must assume that the analyst can delimit a unified, exclusive set of fluctuating language variants that shows meaningful patterns of co-variation with an independent set of variables. Though sensitive to traditional structural categories and relationships in language, the linguistic variable is not necessarily beholden to them, and there are clearly cases where the imposition of a traditional linguistic boundary may actually detract from rather than enhance the understanding of language variation in its social context (Wolfram 1993). Such considerations may be one of the reasons why it might be argued that the linguistic variable is a uniquely sociolinguistic construct that conveniently—if not always comfortably—covers the locus of co-variation between linguistic variants and other variables that include social, historical, psychological, and linguistic factors.

SYSTEMATIC VARIABILITY

As noted, the significance of the linguistic variable lies in its co-variation with other factors, or independent variables. These factors do not typically correlate with the categorical use of a variant but with the relative frequency of variant occurrence. For example, in most varieties of English, there is fluctuation in the production of the final segment of words such as *swimming* and *walking* with a coronal nasal [n] vs. a velar nasal [ŋ], commonly represented in spelling as *-in'* vs. *-ing*, respectively. Studies of the fluctuation between these variants over a half century now (e.g. Fischer 1958, Shuy, Wolfram, & Riley 1967; Kiesling 1996) have shown that the relative frequency of the *in'* variant correlates with various types of social, psychological, and linguistic factors. For example, Fischer (1958) showed that boys tended to use more *-in'* than girls and that differences also related to speaker personality type (aggressive vs. cooperative) and mood (tense vs. relaxed); Shuy et al. (1967) showed that *-in'* was more frequently used by lower status groups than higher status ones, and Kiesling (1996) showed that a group of fraternity men used *-in'* more frequently to project different types of masculine images during fraternity meetings associated with relations of power and solidarity. At the same time, studies have shown that there is variation in the use of the *-in'* variation based on the lexical category of the word, as verbs are more likely to occur with *in'* than nouns (Fischer 1958). In all of these cases, the difference in *-in'* usage is a matter of relative frequency rather than categorical predictability—a

tendency in which variants have a greater or lesser likelihood of occurring under certain conditions.

Factors that correlate with higher and lower frequency levels of a given variant are referred to as *CONSTRAINTS ON VARIABILITY*, where the term “constraint” is used to refer to a factor that systematically correlates with increased likelihood that a given variant will occur. This use is somewhat different from how the term is used in theoretical linguistics, where constraint is used to refer to a condition that characterizes a universal principle of language or a condition that restricts a language-specific rule. Factors that correlate with the increased frequency of a variant are said to *FAVOR* the occurrence of a variant whereas those that correlate with reduced frequency *DISFAVOR* or *INHIBIT* the occurrence of the variant.

Independent variables that co-vary with systematic differences in the relative frequency of a variant are of two primary types, structural linguistic factors related to the linguistic system itself, so-called *INTERNAL CONSTRAINTS*, and social or sociopsychological factors of various types that exist apart from the linguistic system, so-called *EXTERNAL CONSTRAINTS*. Independent linguistic constraints seem to align closely with traditional types of structural units considered relevant in linguistic description. Thus, for systematic phonological variation, the feature composition of the variant (e.g. voicing, sonorancy), phonetic environment (e.g. preceding and following segments, stress patterns), hierarchical status (e.g. syllable position), and grammatical status (e.g. type of morpheme) may be factors that constrain variability. There may also be other factors, such as the lexical condition that high-frequency words favor a variable process over low-frequency words (Myers & Guy 1997). For morphological and syntactic variables, lexical category (e.g. noun vs. verb), phrasal composition (e.g. NP vs. VP, heavy vs. light phrases), co-occurrence relations (e.g. concord, phrasal complements), embedding (matrix vs. embedded clause), and adjacency conditions (proximate vs. distal) may be relevant factors affecting the relative usage of fluctuating variants. These are, of course, the kinds factors considered in traditional qualitative syntactic descriptions. As with phonological variables, this typology may be extended somewhat, so that factors such as the serialization of structures in a narrative sequence, for example, may affect variability along with the reified, sentence-level grammatical categories and constituents generally recognized in syntactic description.

The types of external variables that correlate with the relative frequency of fluctuating variants may include traditional demographic variables (e.g. age, social class, region), constructed social groupings and practices of various types (e.g. communities of practice, social networks), interactional dynamics (e.g. power relations, solidarity), and even personal presentation styles and registers (e.g. performance, mimicking). Although there seems to be somewhat of a disjunction between systemically based linguistic factors and a broadly ranging array of social and psychological factors, most studies of systematic variation include both independent linguistic and social factors within the same description of linguistic variability. Furthermore, one of the principles that has guided variation analysis over the past few decades is the “principle of multiple causes” (Bailey 2002:118), which holds that no single contextual factor can satisfactorily describe the variability observed in natural language. Thus, linguistic factors such as structural composition and linguistic environment may combine with a set of social factors such as age, status, situational context, and so forth in the description of systematic variability.

Though an assortment of linguistic and social variables need to be considered in accounting for systematic variability, not all are of equal magnitude in their orderly effect on the occurrence of variants. To sort out the relative influence of different constraints on variability,

variationists often use a type of multivariate regression analysis model designed by the mathematician and linguist, David Sankoff, known as VARBRUL (Cedergren & Sankoff 1974). Although a variety of commercially available multivariate regression analyses might be applied to determine the significance of various factors on the co-variation of linguistics variants, VARBRUL was designed specifically to handle the kinds of data found in language variation, where not all of the logical cross-products may be manifested and there is often quite unequal distribution of tokens in different cross-product cells in the sample.

VARBRUL is a probabilistic-based statistical procedure that shows the relative contributions of various factor groups to the overall variability of items. The weighting values range from 0 to 1 so that a value of greater than 0.5 in a binomial application indicates that the factor being considered has a favoring effect on the occurrence of the variant, while a value of less than 0.5 indicates a disfavoring effect, whereas in a trinomial application, a value of more than 0.33 would indicate a favoring effect and a value less than 0.33 would be a disfavoring effect. The higher the VARBRUL weighting is, the stronger the effect of the factor is in accounting for the variability. Both Microsoft and MacIntosh versions of this program are readily available and commonly used by researchers in language variation studies. In table 1, taken from Mallinson & Wolfram (2002), results from a VARBRUL analysis of the incidence of syllable-coda, stop-final cluster reduction in a remote Appalachian region of North Carolina (USA) are presented, following a conventional procedure for reporting such data. In this binomial analysis based on the application or non-application of cluster reduction, the input probability represents the overall application mean of cluster reduction apart from the constraining effects. Two independent linguistic variables, or FACTOR GROUPS, are considered in this simplified presentation, the morphemic status of the cluster (e.g. monomorphemic as in *guest* or *mist* vs. bimorphemic as in *guessed* or *missed*) and the following phonetic environment (e.g. consonant as in *guest parking*, vowel as in *guest appearance*, and pause as in *guest //*), and one social factor, ethnic group membership (European American vs. African American). The total Chi square and the Chi square score per cell (i.e. preconsonantal, monomorphemic, African American/prevocalic, bimorphemic, European American/ etc.) indicate the goodness of fit for the quantitative data.

Table 1. Illustration of results of VARBRUL analysis (from Mallinson & Wolfram 2002)

VARBRUL Results
Input probability = .29
Ethnicity
African Americans = .65; European Americans = .33
Grammatical Status
Monomorphemic = .56; bimorphemic = .39
Following Environment
consonant = .80; pause = .37; vowel = .24
<u>Total Chi square = 3.085 Chi sq. per cell = .257</u>

The data indicate that a following consonant (.84) strongly favors cluster reduction over a following pause (.37) and vowel (.24), and that cluster reduction is favored for monomorphemic (.56) over bimorphemic (.39) clusters. Though both following phonetic environment and morphemic status are significant constraints on variability in this study, the following phonetic

environment is a stronger constraint than the morphemic status of the cluster. The analysis also reveals that ethnic group membership is a strong social effect on variability, with African Americans (.65) favoring cluster reduction over European Americans (.35).

Though VARBRUL has become the regular statistical program for determining the systematic effect of various constraints on variability, it is not without limitations. Some of these are a product of the procedural assumptions but some arise from deeper questions about the nature of variation itself. For example, one of the assumptions of VARBRUL is the independence of factor groups, for example, the assumption that the effect of following phonetic environment operates independently from the morphemic status of the cluster. As Sigley (2003) points out, however, both linguistic and social factors are often interactive. With respect to social factors, for example, ethnic group membership may work in tandem with the social factors of age and social class rather than operate independently; similarly, for linguistic factors a suprasegmental effect such as stress may interact with a segmental effect related to the following or preceding phonetic environment. The assumption of independence in effects is therefore questionable, leading to the development of more refined statistical procedures to tease out these interaction effects (Sigley 2003). There are also lingering questions about how such linguistic and social constraints might be represented in a grammar that includes systematic variability.

Another issue raised by procedures such as VARBRUL relates to group and individual variation. VARBRUL analyses are typically based on aggregate data, that is, pooled data for a designated group of speakers without reference to an individual speaker. This procedure obscures variation patterns for individuals, raising the issue of whether or not we can assume that the individual and the group are the same with respect to systematic variation—a kind of homogeneity assumption. Empirical investigations of individual speakers in relation to group norms show that there is impressive regularity from speaker to speaker with respect to the constraint effects on variation, once a reasonable number of tokens are extracted from individual speakers. This leads Guy (1980: 12) to conclude that group norms “recapitulate the generally uniform norms of individuals.” Some of the replicability of constraint effects for individual speakers may, however, be due to systemic principles inherent within language itself. It makes phonetic sense, for example, that syllable-coda cluster reduction would be much more likely take place when the following segment is a consonant as opposed to a vowel, given the phonetic complexity of consonant sequencing without intervening vowels. Other constraints, however, seem language specific, such as the relationship between a following pause and following vowel in their effect on variability (Labov 1994). Though there is sometimes an impressive replicability in terms of constraint effects from speaker to speaker, there may be considerable variation in the overall levels of variation for individual speakers, and some differences in relationships between constraint effects for individual speakers within a group.

Detailed studies of groups and individuals show that one cannot assume that they are always the same with respect to variation, and that there may be considerable variability among speakers within a relatively homogeneous social group. In fact, Dorian (1994) suggests that there is a kind of “personal-pattern variation” that does not co-vary with external social factors (Dorian 1994). Speakers are both individuals with idiosyncratic life histories as well as affiliated members of a complex array of social groups, making it impossible to explain away all individual-based variation in terms of group norms.

VARIATION AND LANGUAGE CHANGE

One of the most fruitful areas for the application of language variation analysis is language change, given the fact that change necessarily involves variation. Speakers do not go to bed one night using a particular form only to wake up the next morning to find the form categorically replaced by another one. Variation exists whether the change involves a gradual, imperceptible shift in the phonetic value of a vowel within a continuum of phonetic space or is an abrupt, readily transparent change involving a major syntactic realignment of phrasal constituents. In the progression of language change, there is a transitional period of co-variation between old and new variants. Though all change seems to involve variation, this does not mean that all variation necessarily implies change. Some variation may be stable, a product of internal systemic and natural performance phenomena rather a reflection of dynamic directionality. Sorting out dynamic and stable variation is, however, not always obvious, and can not necessarily be determined a priori. In most cases, decisions about transitional and stable variation can only be made after examining patterns of variation across time and by considering underlying psycholinguistic and sociolinguistic principles that might help determine the difference.

Although the variability inherent in change has sometimes been ignored in historical linguistics under the traditional assumption that language change could not be directly observed, the empirical analysis of language change at different points in real time and in “apparent time” (i.e. the assumption that differences across different generations of speakers at a given point in time will mirror actual diachronic change) indicates that the variable transition of variants proceeds in an orderly fashion over time and space.

Various models have been proposed for capturing the systematic heterogeneity associated with change. One of the earlier models (Bailey 1973) proposed that change started variably in a limited, natural linguistic environment and then spread from that point to other environments. Furthermore, this model assumed that that variable change in the earliest environment would ideally show greater variability than changes in other environments where the change started later, so that relationships of “more” and “less” with respect to new forms might reflect chronological relations of “earlier” and “later.” This relationship can be illustrated by setting up a simplified, ideal model of the stages of change as in table 2, following Bailey (1973). The table models the progression from the categorical use of one form, A, to another, B, in two different linguistic environments, E₁ and E₂.

Table 2 Variation model of change following Bailey (1973)

/	E₁	E₂
<i>Stage of Change</i>		
Categorical status, before undergoing change	A	A
Early stage, begins variability in restricted environment	A/B	A
Change in full progress, greater use of new variant in E1 where change first initiated	A/B	A/B
Change progresses towards completion with categorically on new variant first in E1 where change initiated	B	A/B

Completed change, new variant categorical

B

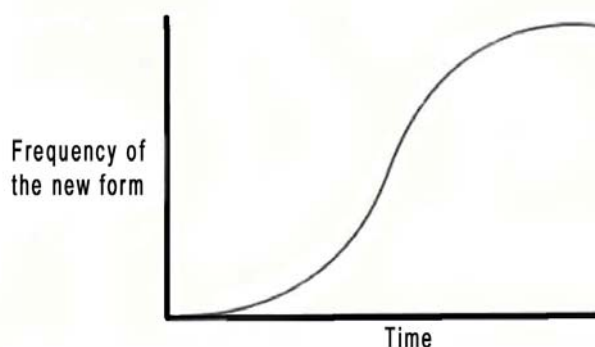
B

The model in table 2 illustrates the “sequential actuation” of change (Pintzuk 2003: 513), where a change starts in one linguistic environment, a more favored environment, and then spreads to a less favored one. However, speakers may also initiate a new form at the same time in different environments, so-called “simultaneous actuation”, or initiate change at different frequency levels in different environments. Cases of simultaneous change, along with other factors that may accelerate or impede a change over time, call into question the assumption that “more” and “less” can be interpreted without qualification as “earlier” and “later.”

The progression of variable change over time has now been subjected to considerable scrutiny in historical linguistics, particularly with reference to syntactic change (e.g. Kroch 1989; Kroch & Taylor 2000; Pintzuk 2003). Though it was originally suggested (Bailey 1973) that more favorable contexts for change might progress at an accelerated rates by comparison with less favorable contexts, research by Kroch and his colleagues supports the conclusion that there is a Constant Rate Effect (CRE) in change, in which the rate of change for each context is the same even though the frequency of competing forms may vary across linguistic contexts. It is important to understand that the label “constant” does not refer to a constant rate of increase in frequency over time, but to the fact that change will be the same across different linguistic environments, so that the progression of change in E_1 , E_2 , and E_n will follow the same trajectory slope.

The study of language change also indicates that there is a prototypical progression slope for variation in time, where change typically follows that of an S-shaped curve (Bailey 1973; Labov 1994; Pintzuk 2003). That is, the replacement of form A with B occurs slowly at the beginning of a change, accelerates rapidly during the mid-course, and then tails off slowly in the final stages of the change, as illustrated in figure 1.

Figure 1. Model of S curve in language change



This model has implications on several different levels. For example, the relatively rapid rate of progression through the mid-course of change sometimes makes this period less accessible to direct observation than the slower periods at the beginning and end points. It may also have implications for the regularity of a change. For example, the role of lexical conditioning in phonological change may be more prominent at the incipient and cessation stages of the change, with lexical diffusion playing a more prominent role at the endpoints. Furthermore, irregularity in constraint effects on variability may be more prone to take place at the beginning and end points of the change than they might be during the mid-course of a change.

The examination of systematic language variation over time offers a unique opportunity to observe language change in progress. More importantly, it has moved historical linguistics away from a long-held belief that diachronic linguistics was limited to post-hoc analysis of the end-products of language change (Weinreich, Labov, & Herzog 1968). Indeed, language change has proven to be one of the most productive venues for studying the nature of systematic variability as well as providing an important proving ground for examining the empirical validity of different models of language change.

VARIATION AND LANGUAGE KNOWLEDGE

As a methodological heuristic, there is little question about the usefulness of variation analysis. Its implications for the representation of speakers' language capabilities and the role of variation in a language grammar, however, is much more disputable. In fact, most theoretical linguists would simply dismiss systematic variability as an artifact of performance with little relevance for a model of language competence. Notwithstanding a few exceptions (e.g. Guy 2003; Anttila 2003), far too little attention has been paid to the role of systematic variability in a grammar of language.

Though the question of speaker knowledge and language variation has been raised on occasion (Cedergren & Sankoff 1974), surprisingly little discussion has focused on the capabilities that speakers might have with respect to language variation as a putative attribute of language competence. Possible claims about speaker capabilities may include the following, ranging from the weakest to the strongest claim.

- Speakers can identify optional (variable) variants
- Speakers can identify the factors that favor and inhibit variable application
- Speakers can identify the relative strength of different constraints on variable application
- Speakers can identify a probabilistic mechanism that generates differential constraint effects
- Speakers can identify frequency levels of variation

There is little psycholinguistic experimentation to support any of these claims, though there is evidentiary documentation that the speaker-hearer can at least recognize optional variants and distinctions between groups and individual speakers based on relative frequency levels. Insight into the finer details of speaker knowledge, such as evidence for knowledge about the relative strength of constraints, the recognition of a probabilistic mechanism for generating constraint effects, and knowledge of frequency levels, however, still awaits psycholinguistic experimentation.

In the early days of variation studies, researchers (Labov 1969; Cedergren & Sankoff 1974) proposed that variation was inherent within a language system and part of speaker competence, and hence should be integrated into the grammar. This was expressed formally in so-called VARIABLE RULES, which incorporated ordered constraint effects, and even probabilities, into the formal generative-style rewrite rules in the tradition of Chomsky and Halle (1968). As formal grammars shifted towards the formulation of universal principles rather than specific language rules, however, the variable rule was abandoned (Fasold 1991). At the same time, reformulated models of grammar raised new issues about the assumed inherency of variation that was once considered to be the cornerstone of variation theory. In the Principles and Parameters model of syntax, for example, it is assumed that parameters are set for a given language in one way or another, while in Optimality Theory in phonology variability is reduced to different rankings among the universal constraints on phonological structure. Accordingly, variation

consists of alternative settings of the parameters or rankings and speakers have at least two grammars, one set in one way and one set in another. As Fasold (2003: 232) puts it “variation in this tradition is not distinguished from bilingualism; such speakers have two more nearly identical grammars and can produce utterances reflecting either or any of them.” Thus, in current studies of variation in historical syntax, variation is now interpreted as grammatical competition between two distinct options that do not ordinarily allow optionality—for example, head-initial versus head-final phrasal structure (Pintzuk, 2003)—so that options correspond to contradictory settings of the parameters.

Though it may be possible to view some kinds of syntactic variation as competition between different grammars of language rather than inherent variability, this interpretation is less comfortably applied across different levels of language. For example, the phonetic realization of vowels may naturally exist on a continuum in phonetic space. Furthermore, given the range of variation in phonology and phonetics, there is an exponential proliferation of possible grammars—into the hundreds and thousands—that seems counterintuitive. The interpretation of variation as a product of competing grammars must also be reconciled with the intuitions of speakers themselves, who may, for example, view variants to exist within a unitary system. Although attempts have been made to reconcile the systematic nature of variability and alternative rankings in Optimality Theory (Guy 2003; Anttila 2003), the mechanisms for incorporating such variability into the grammar still seem somewhat ad hoc, and Guy observes that Optimality Theory seems unable to capture the stability of constraints in the variationist model (Guy 2003: 383).

The potential role of variation in a formal grammar of language remains one of the most challenging—and intriguing—prospects for those who have empirically demonstrated the validity and reliability of systematic variation in language. The systematic nature of structured heterogeneity seems to be indisputable, but its role within a model of language claiming to represent speaker capabilities persists as an issue for variationists and formal theoreticians to contemplate, preferably together.

References

- Anttila, A. (2002) Variation and linguistic theory. In Chambers, Trudgill & Schilling-Estes, 206-243.
- Bailey, C-J N. (1973) *Variation and linguistic theory*. Washington, DC: Center for Applied Linguistics.
- Bayley, R. (2002) The quantitative paradigm. In Chambers, Trudgill & Schilling-Estes, 117-142.
- Cedegren, H. & Sankoff, D. (1974) Variable rules: Performance as a statistical reflection of competence. *Language* 50, 233-55.
- Chambers, J.K. (2003) *Sociolinguistic theory* (2nd edn.). Malden/Oxford: Blackwell.
- Chambers, J.K., Trudgill, P. & Schilling-Estes, N. (eds.) (2002) *The handbook of language variation and change*. Malden/Oxford: Blackwell.
- Chomsky, N. & Halle, M. (1968) *The sound pattern of English*. New York: Harper & Row.
- Crystal, D. (2003) *A dictionary of linguistics & phonetics*. Malden/Oxford: Blackwell.
- Dorian, N. (1994) Varieties of variation in a very small place: Social homogeneity, prestige norms and linguistic variation. *Language* 70, 631-96.
- Fasold, R.W. (1991) The quiet demise of variable rules. *American Speech* 66, 3-21.

- Fasold, R.W. (2003) Language variation in variation and formal syntax. In Preston, D. R. (ed.) *Needed research in American dialects*. Publication of the American Dialect Society No. 88. Durham. Duke University Press. 231-255.
- Fischer, J. L. (1958) Social influences on the choice of a linguistic variant. *Word* 14, 47-56.
- Guy, G. R. (1980) Variation in the group and the individual: The case of final stop deletion. In Labov, W. (ed.) *Locating language in time and space*. New York: Academic Press. 1-36.
- Guy, G. R. (2003) Variationist approaches to phonological change. In Joseph, B.D. & Janda, R.D. (eds.) *The handbook of historical linguistics*. Malden/Oxford: Blackwell. 369-400
- Kiesling, S. F. (1996) *Language, gender, and power in fraternity men's discourse*. Ph.D. dissertation, Georgetown University.
- Kroch, A. S. (1989) Reflexes of grammar in patterns of language change. *Language Variation and Change* 1, 199-244.
- Kroch, A. S. & Taylor, A. (2000) Statistics and grammar: A study of Middle English phrase structure. In Pintzuk, S., Tsouolas, G. & Warner, A. (eds.) *Diachronic syntax: Models and mechanisms*. Oxford: Oxford University Press. 132-163.
- Labov, W. (1966) *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- Labov, W. (1969) Contraction, deletion and inherent variability of the English copula. *Language* 45, 715-762.
- Labov, W. (1994) *Principles of linguistic change*. Vol. I: *Internal factors*. Malden/Oxford: Blackwell.
- Labov, W. (2001) *Principles of linguistic change*. Vol. II: *Social factors*. Malden/Oxford: Blackwell.
- Lavandera, B. R. (1978) Where does the sociolinguistic variable stop? *Language in Society* 7, 171-182.
- Mallinson, C. & Wolfram, W. (2002) Dialect accommodation in a bi-ethnic mountain enclave community: More evidence on the development of African American Vernacular English. *Language in Society* 31, 743-775.
- Myers, J. & Guy, G.R. (1997) Frequency effects in variable lexical phonology. In Boberg, C. Meyerehoff, M. & Strassel, S. (eds.), *Penn Working Papers in Linguistics* 4(1), 215-227.
- Pintzuk, S. (2003) Variationist approaches to syntactic variation. In Joseph, B.D. & Janda, R. D. (eds.) *The handbook of historical linguistics*. Malden/Oxford: Blackwell. 509-528.
- Sankoff, G. & Thibault, P. (1980) L'alternance entre les auxiliaries *avoir* et *être* en français parlé à Montréal. *Langue Française* 34, 81-108.
- Sapir, E. (1921) *Language: An introduction to the study of speech*. New York: Harcourt, Brace, and Co.
- Shuy, R. W., Wolfram, W., & Riley, W.K. (1967) *Linguistic correlates of social stratification in Detroit speech*. USOE Final Report No. 6-1347.
- Sigley, Robert. 2003. The importance of interaction effects. *Language Variation and Change* 15: 227-53.
- Weiner, E.J. & Labov, W. (1983) Constraints on the agentless passive. *Journal of Linguistics* 19, 29-58.
- Weinreich, U., Labov, W. & Herzog, M.I. (1968) Empirical foundations for a theory of language change. In Lehmann, W. P. & Yakov, M. (eds) *Directions for historical linguistics: A symposium*. Austin: University of Austin Press. 95-195.

Wolfram, W. (1993) Identifying and interpreting variables. In Preston, D.R. (ed.), *American dialect research*. Philadelphia/Amsterdam: John Benjamins. 193-221.

Key words: apparent time, constant rate effect (CRE), language variation, constraints on variability (internal, external), linguistic variable, personal pattern variation, S-curve, semantic equivalency, sequential actuation, simultaneous actuation, systematic variability, variable rules, variants (of the variable), variationist, VARBRUL.

Cross-section references:

Dialects, Nancy Dorian, William Labov, syntactic change, phonological change in Optimality Theory, phonological change, copula variation, Constant Rate Hypothesis, final consonant deletion, IN/ING variable, inherent variability, microparametric variation, statistical testing of variation, syntactic variation, variationist sociolinguistics-syntax

BRIEF BIOGRAPHY

Walt Wolfram is William C. Friday Distinguished Professor of Linguistics at North Carolina State University, where he also directs the North Carolina Language and Life Project. He has pioneered research on social and ethnic dialects since the 1960s, authoring or co-authoring 17 books and more than 250 articles on social and ethnic dialects of American English. His current research involves historically and culturally isolated ethnic dialect communities. Professor Wolfram is also vitally concerned with the application of sociolinguistic information to social and educational problems and the dissemination of knowledge about dialects to the public. Wolfram has served as President of the Linguistic Society of America, the American Dialect Society, and the SouthEastern Conference on Linguistics.