

Sociolinguistics and Formal Linguistics

Gregory R. Guy

18.1 INTRODUCTION

Sociolinguistics is often seen as having little relevance to formal linguistic theory. Indeed, a formal linguist once told the present author that sociolinguistics ‘is not linguistics’. While such a comment reveals a narrow and sectarian view of linguistics, it also shows that the relationship between social aspects of language and formal models of language structure is not self-evident – at least, given prevailing theoretical frameworks in the field. So why does this volume have a chapter addressing this relationship?

The most obvious connection lies in the fundamental scientific concern with theoretical adequacy. Chomsky (1964: 29) identifies the lowest level of adequacy as ‘observational’ – providing a faithful representation of the data. Formal linguistic theories that ignore social diversity – a demonstrable characteristic of all languages – do not meet this most elementary criterion, let alone the highly valued descriptive or explanatory adequacy they aspire to. Formalists have glossed over this glaring lacuna by addressing their models to ever-narrower social universes, ultimately the ‘idiolect’ – the grammar of a hypothetical monostylistic invariant individual. But observation denies even this refuge, demonstrating that such an object does not exist in the world. And even if it did, a formal model that accounted for only this would be limited to production, leaving the other two pillars of the language tripod – perception and acquisition – unaddressed, since human beings encounter sociolinguistic diversity in the speech of others, and must be able to learn language and understand others even when they speak differently. Hence the obvious conclusion is that the human language faculty acquires, generates, recognizes and interprets variability, and that formal

theories of this faculty are inadequate unless they account for this.

One contribution of sociolinguistics to formal linguistics is therefore observational and descriptive – it turns up facts that theory must account for. Formal models sometimes flee from the seemingly disorganized multitude of such facts and declare themselves not responsible for the mess. This is an appropriate initial scientific response to complexity – focusing on a subset of facts, idealizing them and ignoring interactions with other facts. But when disciplines mature to the point of achieving some mastery over their subject matter, the gains from idealization are outweighed by the concomitant limits on the theory’s capacity to achieve a more accurate account of reality. At this point a serious science is obliged to engage with what was previously ignored. Linguistics has now reached this phase, and formal linguistics has begun to recognize the need for a sounder empirical basis, and to seek broader empirical testing of theory, which can only be achieved in the messy world of language use. Sociolinguistic studies thus provide a proving ground for theory that is unavailable in the introspective, intuitive paradigm of theoretical research.

Sociolinguistic research also points towards what is needed in linguistic theory to achieve higher levels of adequacy. Adequate formal models must eventually account for variability – an individual’s capacity for stylistic variation, for accommodation, for performing acts of identity by linguistic means – But what are the limits and the landscape of that variability? An adequate formal theory must incorporate social information – social interpretations of variables and ways of speaking; But what kinds of social information are linguistically relevant? Sociolinguistic work is

developing the outlines of the necessary elements of an adequate theory of language – one that accounts not only for narrow Chomskyan competence but also, following Hymes (1972; see Johnstone and Marcellino, this volume), for speakers' sociolinguistic or communicative competence, their capacity to speak and understand in a diverse and varying linguistic world. Encouragingly, this work is meeting complementary developments from the formal side, in theoretical frameworks such as Optimality Theory (OT) and usage-based phonology.

The other side of this relationship is the importance of formal linguistics to sociolinguistics. Sociolinguists have, unfortunately, sometimes adopted a reciprocally sceptical view of the relevance of formal theory to their subdiscipline. Ironically, their reasons for doing so are similar to formalists' reasons for neglecting sociolinguistic results: the models of formalists are diverse, messy and constantly in flux. The impression that outsiders sometimes have of formal linguistics is reminiscent of T. S. Eliot's famous poem, *Prufrock*: 'In the field, the theories come and go, talking of Michelangelo; there is time for a hundred visions and revisions before the taking of the toast and tea'. If theories are ephemeral, focus on irrelevancies, and change a hundred times before tea, an empirical researcher might be wise to avoid making them a central concern of a study that might take years to complete.

Despite sometimes holding such attitudes, sociolinguists cannot avoid theory when they analyse data. Whenever they formulate a hypothesis, or even identify a variable, they are relying on some model of events, some 'theory' (perhaps implicit) of the structure of language. When Labov famously studied coda /r/ in New York City (1966), he made a formal claim to the effect that /r/-less and /r/-ful pronunciations of words like *source*, *car*, *beer* alternated in the speech of New Yorkers, and that the alternants counted as lexically equivalent even though they were phonologically and socially differentiated. This analysis rested on theoretical constructs such as phoneme, word, equivalence, difference and phonological alternation. In general, theoretical models suggest to the sociolinguist what things might be interesting to investigate. A great deal of sociolinguistic research has been devoted to investigating and testing hypotheses derived from theoretical models.

And much sociolinguistic work seriously engages with the construction of formal models. Bailey (1973) proposed a contextually-driven theory of variation and change, in which change is hypothesized to begin earliest and proceed fastest in maximally favourable contexts. Kroch (1989b) tested this model against data on syntactic change and disproved it, proposing instead an alternative

theory – the Constant Rate Hypothesis – that states that changes move at the same rate in all contexts, and are therefore not driven by contextual effects. Guy (1991, 1992) applied the Lexical Phonology model of Kiparsky (1982a) to variable coronal stop deletion (CSD) in English, deriving predictions of an exponential relationship among retention and deletion rates in three different morphological categories, which has been confirmed in studies of several English dialects (cf. Santa Ana, 1992; Bayley, 1994).

Some of the formal models in sociolinguistic work have pioneered in areas where other formal theories have feared to tread. Sociolinguists have been at the forefront of work on interlingual issues, notably the formal models of code-switching proposed by scholars such as Poplack (1980), Myers-Scotton (1993) and Nishimura (1997). Formalists have done little work on linguistic structure above the level of the sentence, but sociolinguists have proposed formal models of discourse level phenomena such as narrative structure (Labov and Waletzky, 1967), conversational turn-taking (Sacks et al., 1974) and therapeutic discourse (Labov and Fanshel, 1977). Thus the relationship of sociolinguistics to formal theory comes full circle, with sociolinguists not only following up the analyses suggested by formal models but also testing and revising them, and proposing new ones.

Of course, as this volume makes clear, sociolinguistics has become a very broad discipline that encompasses a considerable variety of interests and approaches, and not all of these will have the same level of engagement with formal linguistics. Studies of language policy and planning, or language endangerment, are marginal to the construction of formal models of language, whereas sociolinguistic research that examines detailed contextual constraints on phonological or syntactic variation is intimately involved with analysing linguistic structure. But subdividing sociolinguistic interests into those that engage with formal theory and those that do not is not necessarily simple or clear. Consider work on language contact. The Franco-Norman conquest of England in 1066 is a sociohistorical fact with no direct connection to linguistic theory. But the massive borrowing of French and Latinate words into English that ensued has complicated English phonology in ways of considerable interest to formal linguists. Much theoretical work in phonology addresses problems of this sort, such as how to associate certain phonological patterns with specific subsets of the lexicon. Similarly, code-switching was first studied by sociolinguists interested in the social phenomenon of alternating languages in bilingual communities. But it has attracted much formal theoretical attention, due to the insights it offers

into questions of syntactic structure and linguistic cognition.

Therefore we cannot, a priori, partition sociolinguistics into theoretically relevant and theoretically uninteresting components. All its subdivisions have the two connections to linguistic theory we have identified: all have the descriptive connection, revealing facts that theories may eventually seek to explain, and all rely on hypotheses about language structure that are ultimately derived from theoretical models. Since we assume, in principle, that language cannot be adequately understood in isolation from its social embedding, we assume a general relevance of sociolinguistics to formal theory.

In what follows, this chapter considers the following topics: how theory informs sociolinguistics, how sociolinguistic studies address formal issues, and current developments in formal theory that constitute progress towards an adequate model of language – one that has some possibility of accounting for the sociolinguistic facts. But before addressing these topics, we begin with a brief survey of the history of linguistic theory insofar as it has spoken to, or ignored, sociolinguistic questions.

18.2 PAROLE, PERFORMANCE, PERIPHERY: THE MARGINALIZATION OF SOCIOLINGUISTICS IN FORMAL THEORY

Ever since the Neogrammarian movement revolutionized linguistic theory in the late nineteenth century, the mainstream of formal thinking about language has pushed sociolinguistic questions either to the margins of the discipline, or beyond the pale. The Neogrammarians provided a key idealizing assumption that justified this marginalization: the concept of invariance in linguistic processes. The dominant focus of linguistics in the nineteenth century was historical and comparative, leading to the discovery of numerous phonological correspondences between related languages, such as the ‘Grimm’s Law’ relationship between Germanic and Proto-Indo-European (roughly speaking, Germanic underwent a chain shift of the form: /bh, dh, gh/ > /b, d, g/ > /p, t, k/ > /f, θ, h/). Many exceptions to such correspondences were well known, and an emergent dialectological tradition was documenting extensive variability in lexical development. But despite the empirical evidence of variability, the Neogrammarians argued that correspondences arose from EXCEPTIONLESS sound change. The hypothesis of ‘exceptionlessness’ was

justified by a formal model in which words were composed of phoneme-like units on which sound change operated (Paul, 1978 [1886]). When such a unit changed, all words containing it changed simultaneously. Apparent ‘exceptions’ to these correspondences were postulated to result from processes like borrowing and analogical change that operated after the categorical change occurred. Sociolinguistic and dialectological observations of variability were relegated to the periphery of this research paradigm, by definition; they were noise obscuring the systematic big picture. (See also discussion in Kerswill, this volume.)

The Neogrammarian worldview was motivated in part by the discovery of phonological conditioning. For example, Verner (1978 [1877]) famously discovered a prosodic constraint on Grimm’s Law: voiceless stops following an unstressed vowel in Proto-Indo-European are voiced in Germanic, instead of becoming voiceless fricatives. Such discoveries provided a paradigm for formal analysis in linguistics that persists in some measure to the present day: search for invariant generalizations, ignoring variability present in usage; where persistent exceptions to those generalizations are encountered, search for other invariant sub-generalizations delimited by formally definable contexts. This has been a productive heuristic for linguists, but it has also provided formal linguistics with an *oubliette* for disposing of inconvenient facts: one postulates that any exceptions will be accounted for in the future by some more refined statement of the context. At worst, this is unfalsifiable: all formal processes are exceptionless, and any observable exceptions will be explained by future Verners. For linguists who take a Popperian view of their science, such a hypothesis need not be taken seriously.

The invariant view of linguistic process was adapted to synchronic theory by Saussure, the founder of modern formal linguistics. Saussure (1916) elevated the separation of social and formal elements of language to a theoretical principle via his dichotomy between *langue* and *parole*. For Saussure, *langue* is the abstract, systematic aspect of language, while the social diversity of language use is assigned to *parole*. These have unequal status in linguistic science: Saussure privileges *langue* as the object of study of serious linguistics. Sociolinguistic concerns are consigned to the periphery of Saussurean thought, as the study of *parole*.

Saussure’s dichotomy is recast by Chomsky as **competence** and **performance**, and later, I-language (internalized) and E-language (externalized). Competence/I-language is the mental linguistic system governing production – the mental grammar. It is assumed to follow invariant

principles in the Neogrammarian 'exceptionless' tradition. Performance is the product of the system – actual speech, or any other externalized linguistic output (hence E-language). Mediating between the mental grammar and its products are psychological and physiological factors that do not share the property of invariance; hence, they introduce variability into performance. Chomsky cites factors such as slips of the tongue, memory lapses and performance errors as common flaws in everyday language use, and argues that such flaws are not the product of the mental grammar and do not provide useful evidence about its nature (1965).

The competence/performance distinction has logical elegance, opposing a mental machine to its products. But in the give-and-take of linguistic argumentation, it becomes the neutron bomb of generative theory, removing humans from the battlefield, leaving only naked linguistic structures behind. Sociolinguistics is collateral damage in this battle; it provides no data relevant to the construction of formal theory, because it looks only at performance. Chomsky specifically argues that 'observed use of language ... surely cannot constitute the actual subject matter of linguistics, if this is to be a serious discipline' (1965: 4). Performance is thus an even deeper oubliette for the disposal of empirical counter-examples: data from usage bears unknown deviances from the mental grammar, and therefore cannot contradict theories about that grammar.

Chomskyan thought and its theoretical successors thus continue the century-old tradition of consigning sociolinguistic concerns to the periphery. Linguists interested in the topics addressed in this volume therefore find themselves with just two options for relating to linguistic theory. They can accept that their work addresses 'mere' performance, and is consequently irrelevant to formal theory, or they can problematize the assumptions that lead to this conclusion, most importantly the assumption of invariance and the postulated opposition between competence and performance. One encounters, in sociolinguistics, work that follows both of these paths. One trend eschews attention to theoretical issues – a position consistent with the Chomskyan view that evidence from language use is not, in principle, relevant to formal theory.

Another, perhaps dominant, trend in sociolinguistic research has been to reject the sectarian definition that excludes sociolinguistic concerns and denies their relevance to formal models of competence, and instead proposes alternatives that embrace variability and social diversity. This approach found an influential formulation in the work of Weinreich, Labov and Herzog (1968, henceforth WLH), which enunciates two principles that contradict the assumption of invariance. First is 'inherent variability': since language use is

full of variety, it is reasonable to postulate that linguistic systems – mental grammars – generate, perceive and interpret variability. Since no Chomskyan invariant 'ideal speaker-listener' exists, why postulate such a person as the object that our theories seek to explain? Indeed, WLH go further; they argue that if such invariant speakers did exist, they would be social monsters, unable to adapt to changing circumstances, or accommodate to different interlocutors, or to use language to communicate all the social messages performed by variant means.

The second WLH principle is 'orderly heterogeneity': linguistic variability is not random and unstructured, the product of errors. Rather, it is deeply structured by social and linguistic contexts. Essentially, this counterposes a probabilistic interpretation of orderliness to the categorical interpretation prevailing in formal theory. If system and order are equated to exceptionlessness, all valid generalizations must be categorically true; in quantitative terms, they must apply to 100% of cases. But a probabilistic model allows valid generalizations about values below 100%. Social groups are rarely distinguished by the categorical presence or absence of a particular form, but they are often characterized by particular rates of use; thus, in New York City, middle-class speakers use more coda [r] than working-class speakers. And orderly heterogeneity is also evident in linguistic structure: many sociolinguistic variables show strong probabilistic – but not categorical – conditioning by linguistic contexts. Thus, English final CSD is much more likely preconsonantly (*eas'* *side*) than in prevocalic position (*eas'* *end*). Nobody does this categorically, with 100% deletion before consonants and 0% before vowels, but every speaker's usage shows probabilistic favouring of preconsonantal contexts. This process always has 'exceptions', but it is still systematic and orderly.

The challenge to formal linguistics that WLH present, therefore, is to design theoretical models that accommodate variability and probabilistic orderliness. Formal models that continue to adhere to the assumption of invariance and the marginalization of *parole*/performance cannot account for the results emerging from sociolinguistic research, and run the risk of ultimate irrelevance. But many models are emerging that take up this challenge; some of these are discussed in subsequent sections.

18.3 FORMAL THEORY INFORMS SOCIOLINGUISTICS

The fundamental finding of sociolinguistic research is that variation and diversity permeate language. Individual speakers vary their usage

constantly and adaptively, to construct social identities, demonstrate attitudes and manage social relationships. Collectively, speakers are differentiated by geography and time, and by social dimensions of class, sex, ethnicity, age and linguistic experience. Therefore, the fundamental questions in sociolinguistic research are: What is it about language that varies, and why?

Because of the nature of the field, some answers to the ‘Why?’ questions will be linguistic, and some social and historical. Formal linguistics has great relevance to the former, providing potential explanations of constraints on variation. It will not necessarily contribute to understanding the latter: nothing about the structure of language accounts for the social dominance of Spanish over Mayan in Guatemala, or for the spread of Northern Cities dialect features in American English from larger to smaller urban areas (Callary, 1975). But it is often difficult to distinguish one line of explanation from another, and they may interact (if, for example, the linguistic markedness of forms limits borrowability). Therefore, sociolinguistics must allow the possibility of formal linguistic accounts of any patterns of variation it encounters.

The most important contribution of formal theory may lie in answering the ‘What?’ questions. To make progress in appreciating the workings of sociolinguistic diversity, we require concepts and hypotheses about the structure of language – a ‘theory’ that suggests what things are worth studying and where to look for them. Clearly, our investigations are best served by bringing the best theories available to bear on the issues. If we want to know how speakers differ, or what linguistic items they respond to in formulating social evaluations, we would do well to turn to formal theories to delineate the elements and processes that may be relevant.

As an example of the theoretical contributions to sociolinguistic research, consider the variable process of final coronal stop deletion (CSD) in English, a coda-simplification process with parallels in many other languages. It shows classic

traits of a linguistic variable: it is conditioned by linguistic contexts, stylistically constrained (speakers delete less in more careful styles) and socially stratified (higher-status speakers delete less). But the linguistic constraints affecting it have a curious interaction with the social constraints: some constraints are constant for all speakers of English, while others vary from dialect to dialect. Why is this so?

Let us consider two constraints affecting this process. First, as noted above, CSD is markedly affected by following phonological context, with significantly more deletion in pre-consonantal than prevocalic position. Also, the process is sensitive to morphology: stops representing the *-ed* suffix in regular verbs are deleted less often than those that are part of the root morpheme. Relevant values from several studies are found in Table 18.1.

The effect of following vowel (V) and consonant (C) is, as the table indicates, systematic across all dialects. But, strikingly, the effect of following pause (P) varies considerably from group to group. For some speech communities, pause is the most conservative environment, dis-favouring deletion even more than vowels, while for others it is associated with high deletion rates. The first discovery of this dialect-specific effect of pause was made by Guy (1980), who found that New Yorkers favoured deletion before pause, while Philadelphians disfavoured it. The effect was systematic within each speech community, as Table 18.2 illustrates. Everybody in both communities (except two Philadelphians with small *ns*) has more deletion before C than V, but 18 of 19 Philadelphians show pause as the least favourable environment for deletion (P deletion rates below V), while no New Yorkers have this ranking of vowel and pause.

A similar issue arises with respect to morphology. Most speakers in most dialects systematically show higher deletion rates in monomorphemic (M) words like *mist*, *pact*, *bold*, vs reduced rates in regular past tense verbs (R) like *missed*, *packed*, *bowled*. But speakers vary considerably in their

Table 18.1 Effects of morphological class and following segment on English coronal stop deletion

	Morphological			Following segment		
	M	I	R	C	V	P
Bayley	0.53	<	0.62	0.73	0.27	<0.46
Santa Ana	0.55	>	0.43	0.62	0.33	>0.32
Lim and Guy	0.60	<	0.65	0.61	0.37	<0.48
Guy	0.64	>	0.55	0.66	0.19	<0.37

M, monomorphemic words; I, irregular past tense verbs; R, regular past tense verbs; V, vowel; C, consonant; P, pause.

Source: (from Bayley (1999), Guy (1991) Lim and Guy (2003) and Santa Ana (1992))

AQ: Confirm values are <0.62, <0.43, <0.65, <0.55 for I and <0.46, <0.32, <0.48, <0.37 for P

AQ: is not in References: please supply

Table 18.2 Comparing constraint effects in two speech communities: following context constraint orderings for coronal stop deletion in New York and Philadelphia

	Following context			
	Number of speakers showing the order:			
	C>V	V>C	V>Pause	Pause>V
Philadelphians	16	2	18	1
New Yorkers	4	0	0	4

C, consonant; V, vowel.

treatment of the irregular past tense verbs (I) that take both a root vowel change and a coronal stop affix in the past, such as *lost*, *kept*, *told*. Guy and Boyd (1990) found that the treatment of this morphological class varied with age: older speakers treat irregular verbs conservatively, with lower deletion rates, while younger speakers are more likely to delete them (Figure 18.1).

Such findings provoke an obvious question: Why are some constraints constant across speakers and dialects, while others are inconstant, either dialect-specific or age-graded? Since sociolinguistics is simultaneously concerned with linguistic structure and social context, we might entertain both social and linguistic explanations of these facts. On the social side, a possible hypothesis is that speakers arbitrarily select linguistic items for social-indexical purposes. This is akin to the arbitrariness of the linguistic sign: just as a book does not intrinsically demand the label *book*, since it is labelled *hon* in Japanese and *kitab* in Arabic, so there is no intrinsic reason that pronouncing coda /r/ is socially favoured in New York. Therefore, we might suppose that the prepausal deletion rate for coronal stops is arbitrarily selected as a dialect marker, and the deletion rate in irregular verbs has been arbitrarily associated with particular age cohorts.

Pending a more comprehensive analysis of what variables and constraints have what social interpretations and associations, we cannot rule out this account on purely social grounds. But the fact that numerous studies of CSD show the same effects for many other constraints, leaving just these two contexts to vary greatly across speakers, should give us pause. Perhaps some linguistic property makes these constraints more amenable to social differentiation.

This is where formal theory comes to our aid. Many strands of phonological theory agree that there is a universal hierarchy of syllable types, with CV being the most basic, found in all languages, while other structures, including complex codas, complex onsets and zero onsets, are more marked and typologically less common. These principles are explicitly incorporated in CV

phonology (Clements and Keyser, 1983), and in postulated OT constraints such as NoCoda, *Cx coda, Onset, etc. Hence, these theories suggest a prediction, to the effect that ... VCCCV... sequences, in phrases like *east side*, are universally more marked than ... VCCV... sequences, like *east end*. This being the case, a process that simplified the coda sequence should always be more strongly motivated in the former than the latter. If these are linguistic universals, we should not expect to find any dialect or speaker that contradicts them, which, essentially, is what the data show – everybody deletes more before consonants than before vowels.

But what do formal models say about following pause? Eligible words for English CSD have the form ...VCC, so when no other subsequent segments are considered, they will violate NoCoda and *Cx coda constraints just like eligible words occurring in other contexts, but are they more or less marked than prevocalic or preconsonantal contexts? Formal theories are largely silent on this point; an answer would depend on a formal analysis of non-speech (What are the distinctive features of silence?), or on an additional set of hypotheses about syllabification, stop release, etc., which would not be universal in nature. Therefore, it is a fair conclusion from phonological theory that the effects of prevocalic and preconsonantal contexts are determined by linguistic universals, but the effect of following pause is not universally defined, and hence is available for dialect-specific, dialectally arbitrary treatments.

Formal theory suggests a similarly ambiguous status for the treatment of CSD in irregular verbs. The difference between monomorphemic and regular past tense forms is modelled in several ways in different theories: by functional constraints against loss of information; by a historical constraint against loss of morphological distinctions (Kiparsky, 1982b, the 'distinctness condition'); and by the different derivational histories of the two classes (underived vs derived, cf. Guy, 1991). But all models assign a distinct status to the morphological marker in *missed*, *bowled*

that is absent in *mist*, *bold*, with a consistent prediction that deletion of the past tense forms is disfavoured.

But what about irregular past forms? There is no obvious universal explanation for how they should be treated by a deletion process. The final stops in *left*, *told* mark tense redundantly, so they have no greater functional load than those in monomorphemes. Their derivational histories depend on specific hypotheses about what form classes exist and how they are stored and generated, on which there is little theoretical consensus. Therefore, we can again entertain a reasonable hypothesis that while the difference between monomorphemes and regular past tense forms is subject to universal conditions, the status of the irregular forms is precisely irregular, subject to different analyses and different treatments by the deletion process.

This is what Guy and Boyd (1990) argue. They see the age-grading in Figure 18.1 as a product of successive reanalyses of these words by English speakers as they develop linguistically. The youngest (pre-teen) speakers classify them as strong verbs, lacking a suffix (thus *keep~kep*, analogous with *feed~fed*); hence, final stops are almost completely absent. By adolescence, most speakers construct a special morphological class for such words, combining traits of the strong and weak classes, but without any internal morphological analysis (thus *keep~kept*). Hence, they show the

same deletion rates as monomorphemes, which also lack internal morphology. Finally, in adult life, many speakers proceed to a complex analysis in which the final coronal stop acquires a separate morphemic analysis, no doubt by analogy with the regular affix in *missed*, *bowled* (thus *keep~kep#t*); hence they begin to partake of the morphological resistance to deletion that regular past forms exhibit.

In both of these examples we find that theoretical analysis supplies a principled basis for distinguishing the constant constraints from those that vary dialectally or generationally; some elements of linguistic structure are universal, and cannot vary between speakers, dialects or languages, while others are not universally specified, and hence are available for social differentiation.

This conclusion is supported by other studies of sociolinguistic variation, beyond phonology and beyond English. A case in point is Cameron's (1993) study of subject pronoun expression in Spanish. Spanish is a PRO-drop language – subject pronouns are optional. In the meaning 'I want', a Spanish speaker can say either *Yo quiero* or *Quiero* with the pronoun *Yo* 'I' expressed or omitted. Cameron reports a major dialect difference between Puerto Rico and Spain: Puerto Ricans use many more overt pronouns than their Spanish counterparts. He also identifies linguistic constraints on this process, including switch reference: speakers are more likely to use an overt

AQ: Please underline whether underline should be given.

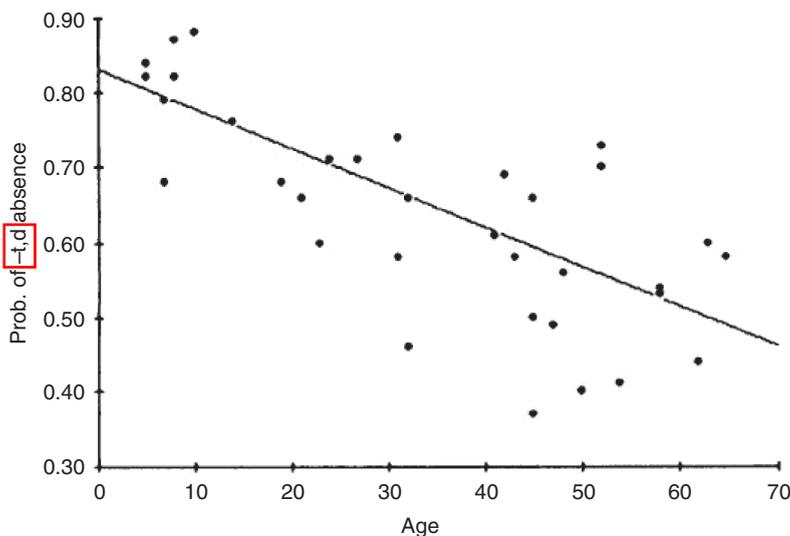


Figure 18.1 Probability of -t,d absence in irregular past-tense verbs, by age (from Guy and Boyd, 1990: 8.)

pronoun to mark a change in subject between the previous clause and the current clause. Most linguistic constraints are common across the two dialects, but one is treated differently: specificity of reference (the difference between references to specific individuals and generic references meaning ‘somebody, one, people in general’). For the second-person singular (2sg) pronoun *tú*, this constraint is significant in both countries, but with opposite effect. Puerto Ricans use overt *tú* more often when making a generic reference, while Madrileños favour overt *tú* for specific reference (Table 18.3).

What does linguistic theory say about these constraints? The switch-reference effect has a clear functional motivation: the low information alternative (with no overt pronoun) serves as a gap onto which earlier references are projected. But the higher information form, with an overt pronoun, is used to signal a new subject. English uses contrastive stress in much the same way. In a sentence like ‘Billy talked to Paul about it, and HE said ...’ the stressed ‘HE’ refers to Paul, as a new subject, whereas an unstressed alternative ‘... and he said ...’, would ordinarily indicate a continuation of the same subject, i.e., Billy. In each case, greater emphasis on the subject (via an overt pronoun or contrastive stress) signals a switched reference. This is a possible universal. But the specificity constraint has no such universal association with overt or zero pronouns. English, which only allows overt pronouns, uses the 2sg pronoun for both meanings. A sentence like ‘You know what it means’ can mean either specifically ‘you (the hearer) know ...’, or generically ‘one knows ...’. This suggests no intrinsic association of overt pronouns with either meaning; rather, such an association is arbitrary, and two Spanish dialects have made different selections for indicating genericity and specificity.

18.4 SOCIOLINGUISTIC CONTRIBUTIONS TO FORMAL THEORY

In formal linguistics one encounters many examples of principles that are originally proposed as

categorical – unsurprising, given the categorical bias of linguistic theory. A case in point is the Obligatory Contour Principle (OCP), a generalization in phonology to the effect that languages prefer alternations (or ‘contours’) over repetition of identical elements. Thus the preferred structure for syllables and words involves alternating consonants and vowels, not strings of one or the other: CVCV over CCC or VVV. Similarly, tone languages prefer alternating tones, not strings of identical tones: HLHL, not HHH or LLL. Indeed, the OCP was first proposed by Leben (1973) to account for categorical tonal alternations in African tone languages.

A common finding of sociolinguistic research, and of other quantitative studies of language, however, is that essentially the same principles are observed in linguistic variation, but in a probabilistic, rather than categorical form. This discovery has been termed the ‘stochastic generalization’ (Bresnan et al., 2001; Clark, 2005): generalizations that may be categorically true in one language exist as quantitative constraints on variation in another language or social variety.

This is the result Guy and Boberg (1997) find for the OCP in their research on English CSD. The segment preceding a deletable final stop has a proven constraining effect on the likelihood of deletion; roughly speaking, obstruents promote deletion more than sonorants. Guy and Boberg accounted for this effect in terms of the similarity between the preceding context and the deletion target (the final coronal stop): greater similarity (measured in terms of shared distinctive features) favours deletion. The values appear in Table 18.4.

The results show that segments that share two of the defining features with [t, d] all promote deletion more than those that share only one feature, while preceding vowels, which are completely dissimilar, almost never trigger deletion. Note further that adjacent identical stops are categorically prohibited in English (*tt, *dd) – itself an OCP effect. Thus the results demonstrate a continuum: the phonology of English disfavors sequences of segments to the extent that those segments are similar; sequences of partially similar segments are partially suppressed by means of CSD, and sequences of completely similar

Table 18.3 Subject pronoun expression in Spanish: dialect differences in the effect of specificity of reference with second-person singular pronoun *tú*

	<i>San Juan, Puerto Rico</i>			<i>Madrid, Spain</i>		
	<i>Per cent overt pronouns</i>	<i>Factor weight</i>	<i>n</i>	<i>Per cent overt pronouns</i>	<i>Factor weight</i>	<i>n</i>
[+specific]	48%	0.51	145	40%	0.72	58
[–specific]	69%	0.72	188	19%	0.50	150

Source: Camerson (1993: 325)

Table 18.4 Preceding segment effect: an OCP analysis

Preceding segment	Deletion		
	N	Per cent	Factor weight
/t, d/ [+cor, -son, -cont]	–	(categorical absence, i.e. 1.00)	
/s, z/ [z] [+cor, -son]	276	49	0.69
/p, b, k, g/ [-son, -cont]	136	37	0.69
/n/ [+cor, -cont]	337	46	0.73
/f, v/ [-son]	45	29	0.55
/l/ [+cor]	182	32	0.45
/m, ŋ/ [-cont]	9	11	0.33
vowels –	–	(nearly categorical retention, i.e. 0.00)	

Source: [Guy and Boberg \(1994\)](#) [corpus: 3 Philadelphians]

(i.e. identical) segments are completely suppressed, banned in underlying forms, and avoided in derived forms by epenthesis (thus, bisyllabic *raided* vs monosyllabic *rained*).

Results of this sort have two-fold implications. First, they confirm the theory; here they show that OCP is an active principle in English. But, secondly, they show the principle to have graded, probabilistic effects, not only categorical ones. In the Guy and Boberg data we see not an ‘obligatory’ principle, but rather a contour preference whose force is inversely proportionate to the slope of the contour: flat contours are prohibited, gentle slopes are disfavoured and steep slopes are promoted. This preference operates gradiently in sociolinguistic variation, not as a ‘knockout’ effect. Words with disfavoured consonant sequences like *act* are not categorically deleted in some social dialects or styles and retained in others; rather, they are variable in all dialects, but always show higher rates of deletion than those with more favourable contours, like *old*. The sociolinguistic findings thus test the formal model – the OCP – and in this case confirm it, but recast it as a stochastic generalization.

The OCP example shows a formal principle, conceived in the Neogrammarian tradition as discrete and invariant, being reformulated in the light of sociolinguistic evidence as a probabilistic generalization. What might the impact of sociolinguistic evidence be on larger elements of formal theory, such as an entire grammar? Formal linguistics since Chomsky attributes a speaker’s capacity to use language to a mental grammar that codifies and encompasses linguistic knowledge. This is constructed by speakers in the course of language acquisition, based on input from the language they hear spoken around them (along with, presumably, some universal elements).

This model presupposes that the speaker is embedded in a social universe, with other speakers from whom the individual hears evidence needed for language acquisition, and with whom the individual communicates using his or her

mental grammar. So an obvious question is, how similar or different are the mental grammars of individuals who speak what we informally call ‘the same language’? Given the diversity that sociolinguistic research documents, even among members of the same speech community, what elements of their mental grammars are diverse, allowing differentiation, and what are the same, allowing them to communicate in a language that they believe they share?

In the mainstream tradition the mental grammar is an apparatus with categorical, invariant properties. Hence even a small alteration to its internal structure, if it produces any difference in output, would count as a discretely different grammar. Linguistic diversity, therefore, would imply that speakers have grammars that differ in some component or set of components (features, rules, constraints, constraint rankings, etc.). Different speech styles would be handled the same way: to shift among different speech styles or registers, speakers would have to switch between discretely different grammars. By implication, styles and speakers would exhibit categorical differences: speaker A would use some form in some context that speaker B did not or could not use.

In the social world, it is clear that diversity is orderly: socially proximal individuals are linguistically proximal as well; family members speak more alike than those who are unrelated; people in the same community or dialect region are more alike than those who live farther away; and people of the same ethnicity, class or age group also tend to be more similar than those of other groups. What does this scale of sociolinguistic similarity correlate with in the mental grammar? Is the scale a gradient continuum, or is it composed of discrete steps, so speakers who are more dissimilar simply have a larger number of quantum differences in their grammars?

Sociolinguistic research suggests nuanced answers to such questions. We encounter both discrete and continuous differences between varieties, and we encounter different orders of

AQ: Please provide data.

AQ: is not in References: please supply

quantitative diversity. First, much sociolinguistic variability is indeed continuous. One finding, by now classic, is that speech communities are replete with variables showing social stratification: higher-status speakers use more and lower-status speakers use less of a given variant and, simultaneously, all speakers use more of that variant in more careful styles. Labov's (1966) findings on the use of coda /r/ in New York City cogently illustrate this pattern.

The data in Figure 18.2 clearly suggest a continuum of usage; New Yorkers make finely graded quantitative adjustments in their individual production, and are collectively finely graded in /r/ usage by socioeconomic status. It is implausible to perceive in these data a series of discretely different grammars, each associated with a particular speaker in a particular speech style, each displaying some quantum difference from socially or stylistically adjacent grammars. Instead, it suggests a quantitative component of the mental linguistic capacity that controls the frequency of coda /r/ – in other words, a probability. Each speaker has a characteristic value for this probability, but they can adjust it continuously for stylistic purposes, to accommodate the social

demands of a situation, or for expressive value. Social proximity between individuals is reflected in their settings of this probability: one's experiences in language acquisition, daily interaction and accommodation, and attitudinal factors such as solidarity, all combine to make it likely that speakers who belong to common networks and who share social characteristics such as class, ethnicity, etc., will have common values for these probabilities. And what the speech community as a whole shares, including New Yorkers of different social classes and ethnicities, is a common social evaluation of the use of the /r/ variant: they all understand it is associated with higher-status speakers and more formal styles, and reflect this shared evaluation in their own usage.

However, beyond these scalar patterns of variation associated with style, class, age, etc., we also encounter discrete differences. The data presented above in Tables 18.2 and 18.3 demonstrate this, showing discrete differences between speech communities. Notice, however, that those examples involved the effects of context on linguistic variation: the effect of a following pause on English CSD, and the effect of specificity of reference on pronoun expression in Spanish. Notice further that

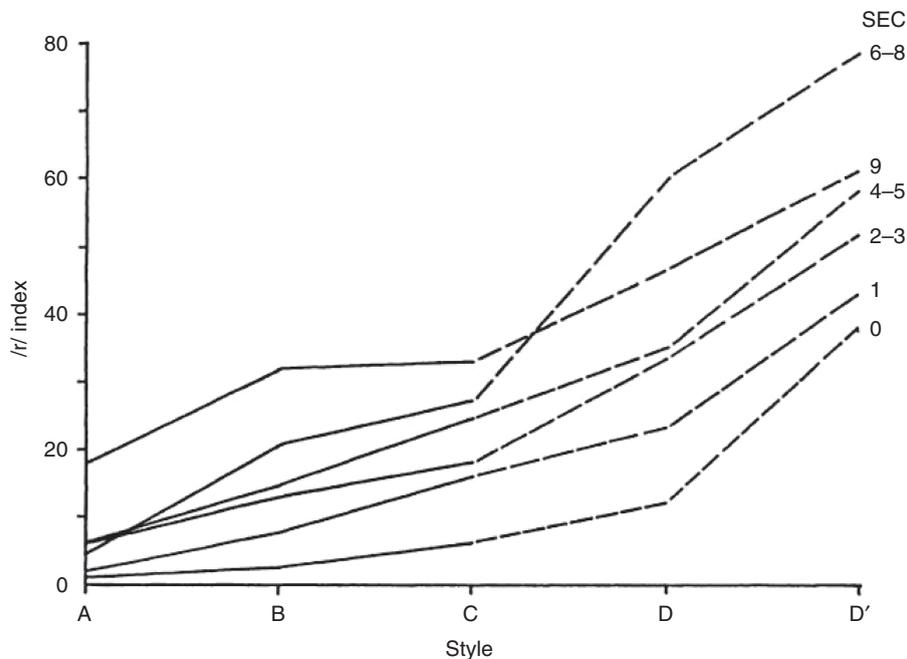


Figure 18.2 Class stratification of coda /r/ in New York City. (from Labov, 1972: 114). Socioeconomic class (SEC) scale: 0–1, lower class; 2–5, working class; 6–8, lower middle class; 9, upper middle class. Style scale: A, casual speech; B, careful speech; C, reading style; D, word lists; D', minimal pairs.

these discrete differences lay between different speech communities – Philadelphia vs New York, San Juan vs Madrid – whereas the continuously varying data of Figure 18.2 come from speakers in the same community, in fact the same neighbourhood (the Lower East Side of New York).

This suggests a basic distinction between two dimensions of quantitative variability. We distinguish overall rates of use of a variant (a context-free statement of the alternation) from differential rates of use by linguistic context (context-sensitivity). Quantitative variation in the overall rate of use of social variables, differentiating classes, speech styles, etc., is in principle continuous, and speakers within a given community are accustomed to, and accomplished at, varying freely in this dimension. But the contextual effects, although they may also be probabilistic, do not vary freely within the grammar of an individual, or within a community. Thus there is no social class of speakers in Philadelphia, for example, for whom pause is a more favourable context for deletion, and no speech style in which Philadelphians use pause as a favourable context. Hence we advance a hypothesis: the rankings, and quite possibly the specific probabilities, associated with constraints on linguistic processes are a fixed feature of the grammar of a speaker, and are one of the shared linguistic characteristics common to all members of a speech community.

This distinction is given formal representation in the ‘variable rule’ model (Labov, 1969; Cedergren and Sankoff, 1974; Sankoff, 1978). This model associates an ‘input’ probability (p_0) with every option or rule in the grammar; in addition, context-sensitive processes may have probabilistic weights associated with relevant contexts (contexts i, j, k , etc., are associated with weights p_i, p_j, p_k). In this model, the value of p_0 is what varies within a community, whether stylistically or socially, while the values of specific constraints p_i, p_j, p_k are fixed. This is an original contribution of sociolinguistic research to the theory of grammars, specifically to the question of grammatical similarity and difference.

18.5 TOWARDS AN ADEQUATE FORMAL LINGUISTICS

What does an adequate theory of language need – one that accounts for speakers’ capacity to produce and deal with diversity? The first requirement is that it must accommodate variation. The Neogrammarian-inspired model of exceptionless formal linguistic processes has been mistaken in much of formal linguistics for a design principle of language; in fact it is a hypothesis, which has been

tested against reality and found wanting. Language is neither Platonic nor digital; it is a neurological, physiological and social system, an analog device that doesn’t always do the same thing in exactly the same way. An adequate formal theory must recognize inherent variability, and generate it.

Secondly, the sociolinguistic evidence suggests that an adequate model must incorporate quantification. This point is still the subject of some debate, as some scholars propose to leave the grammar unquantified and derive quantitative regularities observed in language use by other means, as an epiphenomenal consequence of the range of possibilities (cf. Anttila’s OT analyses of variation – see below), or by some mechanism external to the grammar proper that selects among different grammars. We shall argue, however, that these alternatives are not adequate to account for the full range of quantitatively orderly heterogeneity.

Thirdly, an adequate model must accommodate social information. **It requires representations that model speakers’ capacities to speak in socially adaptive ways, and to perceive the social significance of the usage of others.** Since individuals smoothly vary their usage both expressively and responsively, thereby communicating social meanings and adapting to social expectations, an adequate formal model of this capacity must incorporate a social semantics not only of words but also of linguistic forms and processes. The formal account must also be capable of representing diversity at the community level, differences in the usage of others that speech community members can passively recognize the social significance of, even if their own grammars do not productively generate them.

The sociolinguistic adequacy of a formal model can therefore be evaluated on the basis of these three criteria: accounting for variation, quantification and social information. Let us consider some formal models that have been proposed to address sociolinguistic issues, and how they fare by these criteria.

The first issue, of modelling variation, has been widely addressed, by a variety of formal approaches. The variable rule (VR) model, mentioned above, uses a generative framework in which probabilities are attached to operations; conventional obligatory processes have a probability of 1, but probabilities of less than 1 indicate variable outcomes. This is the model most widely used in sociolinguistic research on variation. In OT, several proposals have been made to generate variability using variable or underspecified constraint rankings (Anttila, 1997, 2002; Nagy and Reynolds, 1997): instead of a fixed constraint hierarchy that always gives the same outcome, speakers are hypothesized to vary among several rankings of the relevant constraints which select

AQ:
Please
check
sentence.

different variants. In treatments of historical change, diachronic variability has been modelled by means of alternation between grammars that generate old and new variants (Kroch, 1989b, 2000; Yang, 2000). All of these models are adequate for generating variable outcomes.

Greater differences between approaches are found when we consider the second criterion, quantification: How should a formal model account for the probabilistic patterns of natural language? Do they result from one grammar or from alternation between grammars? Should they be generated directly, by quantification within the linguistic system, or derivatively, as a secondary consequence of the way a grammar operates to generate a range of options, or the diversity of representations or inputs?

The quantificational strategy that involves the smallest departure from the theoretical tradition of invariance is grammar competition. In this approach, traditional invariant grammar is preserved, and variation and change are taken as evidence that speakers command multiple grammars, each one discretely generating one observed variant. These alternative grammars are selected with different probabilities to produce the quantitative patterns that are found in language use. This approach is exemplified in work on diachronic change by Kroch (1989b), Yang (2000) and others. Kroch (1989a), for example, takes this approach to model the development of periphrastic *do* in English. In early Middle English (ME), before AD 1300, negative constructions inserted *not* after a verb or auxiliary (e.g., *They know not what they do*), and questions were formed by inverting subject and verb (e.g. *Slept you well?*). In late ME and early Modern English, modern equivalents with auxiliary *do* began to appear (*They don't know what they're doing; Did you sleep well?*); for the next 400 years, usage varied between the two forms, with the modern constructions steadily increasing in frequency.

Kroch accounts for these facts in terms of a parametric shift. The grammar of early ME permitted the main verb to raise to an auxiliary (or **INFL**) position, so that it could do whatever auxiliary verbs do (in Modern English other auxiliaries also occur before *not* and invert with subjects to form questions: e.g. *They must not know; Have you slept well?*). But in late ME an alternative grammar emerged that did not permit verb-to-INFL raising, so that when the auxiliary position was separated from the verb by another element, such as an inverted subject or a negative, *do* was inserted as an empty auxiliary to carry tense and agreement. In Kroch's analysis, during the four or five centuries when this variation was evident, speakers entertained two mental grammars, one with and one without V-to-INFL raising, and their

variation consisted of choosing one grammar or the other for each relevant sentence they uttered. What changed across time was the probability of speakers selecting one or the other grammar, while within each grammar, there was no variation.

While providing an elegant treatment of historical change, this strategy does not offer a general quantitative model for sociolinguistic variation. First, it does not, strictly speaking, solve the problem of variation; rather, it relocates it from inside the grammar to outside. To account for the quantitative facts of orderly heterogeneity – all the subtle probabilistic variation evident in sociolinguistic usage – this approach requires an additional theory of grammar selection. Secondly, it runs afoul of Occam's razor – the elementary scientific principle of economy, that proliferation of explanatory elements should be avoided. For the English *do* case, Kroch requires two grammars that are identical in every respect except for the V-to-INFL parameter setting. But at any given moment, languages typically have not one but many variables – thus ME also has variable deletion of *that* complementizers (*I thought THAT he knew vs I thought he knew*), among other variables. If each alternant of each variable requires a separate grammar, and these variables intersect, so that a speaker may in one utterance use periphrastic *do* and complementizer deletion and, in the next, periphrastic *do* but complementizer retention, and so on, then a language with *n* binary variables would require speakers to maintain *n!* (*n* factorial) mental grammars, each trivially different from the next. Besides constituting a massive violation of Occam's razor, this would rapidly exhaust the storage capacity of the brain, if the language had as few as a score of variables.

The alternative to grammar competition is to step off the Neogrammarian path and locate variation inside the grammar. But there are still competing strategies on how to formalize this so as to achieve quantitative adequacy. The VR model uses explicit quantification; in principle, it permits the probabilistic quantification of all grammatical operations and contextual conditions. This allows it to achieve excellent accuracy in accounting for the quantitative patterns observed in sociolinguistic variation. However, some scholars have argued that incorporating a probabilistic component in the grammar gives an overly powerful formal model, capable of generating patterns that do not occur. A notable alternative is found in versions of OT using variably ordered constraints (VOC), which treat quantitative patterns as, in a sense, epiphenomena – statistical reflections of the range of possibilities permitted by grammatical options, rather than the product of a specific probabilistic element. This approach is exemplified by Anttila's treatment of phonological

AQ: need to define INFL? inflection?

variables in several languages (Anttila, 1997; Anttila et al., 2008).

As a simplified example, consider a VOC treatment of English CSD. Suppose there were only three relevant constraints: one disfavouring syllabic codas (NoCoda), one disfavouring complex codas (*CxCoda), and one favouring the faithful articulation of underlying consonants (FAITH). If the ranking of these constraints was unspecified in the grammar, and randomly varied whenever speakers articulated a relevant word, they could occur in any of six possible orders (NoCoda highest, followed by the other two in either order; *CxCoda highest, followed by either order of the other two, and FAITH highest, with either order of the other two). Following standard OT conventions, we assume the grammar selects outputs consistent with (i.e. incurring the fewest violations of) the highest ranked constraints. Therefore, when FAITH is highest-ranked, words like *act*, *mist* are pronounced with final /t/, but when either of the other constraints outrank FAITH, a pronunciation with a deleted /t/ results, because NoCoda and *CxCoda are both violated by candidates preserving the final stop. Since two of the six possible orders generate retention of /t/, and the other four generate deletion, this model predicts deletion should occur in two-thirds of occurrences.

Using the VOC strategy, Anttila has achieved impressively accurate quantitative models of several cases of linguistic variation. He also argues that VOC explains dialect patterning – why certain combinations of variants are dialectally attested, while other patterns we might imagine, but which are not generated by any possible order of constraints, do not occur. These results present a serious challenge to other formal models addressing these questions. However, the approach is inadequate as a general model of sociolinguistic variation. As we have noted, quantitative variation has two aspects: differences in overall rates, which vary between speakers and across speech styles, and differences in constraint rankings and contextual effects. VOC can model the latter, but lacks any formal treatment of the former. The only way predicted probabilities can differ in VOC, among styles, speakers or dialects, is by varying the number of constraints that are variably ranked. In the simplified example above, three constraints have $3! = 6$ possible rankings; if four constraints were involved, they would have $4!$, or 24 possible rankings. So adding more constraints to the set that are variably ranked gives more possible rankings, and therefore different possible proportions among the selected variants, but it also means the variation will now occur or fail to occur at different rates in whatever contexts are affected by the constraints that have been added to the mix.

To illustrate this point, we can extend the previous example. Suppose, in another speech style, or a different group of speakers, a fourth constraint is added to the variably ranked group, an OCP constraint prohibiting sequences of overly-similar consonants, ruling out cases like *act*, *mist*, but permitting *left*, *old*. When this constraint outranks FAITH, deletion is selected in words like *act*. Hence when OCP is variably ranked with respect to the other three, undeleted forms of *act* are selected only one-quarter of the time (the possible orders in which FAITH is highest ranked), rather than one-third, raising the expected deletion rate from 67 per cent to 75 per cent. But this constraint has no effect on deletion of words like *old*, which will continue to show deletion in 67 per cent of all cases. Hence the consequence of involving an additional variably-ranked constraint is to introduce a new difference between contexts, not to change deletion rates across the board. As we have noted, this is not how sociolinguistic variation actually works; styles and social groups in the same speech community differ in overall rates of use of variables, not in contextual constraints.

Finally, consider the third criterion, the formal modelling of social information. The above example shows the difficulties OT encounters in this respect; it could incorporate social information by associating styles or social indexicality with certain constraint rankings, but this incorrectly predicts regular differences between speakers or styles in contextual constraints on variables. VR models treat social information as probabilistic constraints on production; thus stylistic constraints fit into the model in the same way as linguistic constraints. In speech community studies using VR, it is routine practice to treat social groups in the same way; so age or gender groups, for example, are treated as if they were contexts for usage associated with probabilities that increase or decrease the use of a variable. However, this practice has a hazy theoretical status. It may more or less accurately model the collective behaviour of a community, but it is not clear what it means for the grammar or usage of an individual.

An important lacuna in all of these theoretical approaches to sociolinguistic variation is that, since they focus on variable production, they have little to say about the perception and mental representation of diversity, so that speakers' passive knowledge of the social significance of the usage of others is not explicitly accounted for. It is a central argument of sociolinguistics that speakers understand the social significance of the linguistic behaviour of other members of their speech communities, even when these others are doing things with their language that the speaker does not personally do. This understanding is reflected in shared evaluations of variables, and it influences

production in such ways as the shared directions of style shifting. But how is this knowledge represented in the mind? Should it have formal representation in the mental grammar? To conclude our consideration of formal models, we turn now to a theory that puts such representations at the centre of linguistic structure.

18.6 NEW DIRECTIONS: USAGE-BASED GRAMMAR

An important theoretical development that explicitly seeks to offer a formal account of usage, variation and change, and to provide a model of speakers' knowledge that incorporates social and stylistic information, is the emergence of usage-based approaches to grammar, notably Exemplar Theory (Bybee, 2001; Pierrehumbert, 2001). This theory is a radical departure from mainstream formal linguistics; it embraces variability and rejects categorical, exceptionless processes. Indeed, it denies the existence of general abstract processes, along with abstract representations of words and segments. Instead, Exemplar Theory (ET) begins by postulating that speakers remember, in rich phonetic detail, the tokens of words they hear pronounced, or produce themselves. Therefore, speakers have stored in memory the full range of variation they have encountered, and use these memories (the 'exemplar cloud') as targets for their own productions, which necessarily also vary. Words are not decomposed into strings of phonemes, but remembered holistically; hence each word can behave distinctly, showing idiosyncratic patterns of variation and change that diverge from other words containing the same phonemes. The substance of other approaches to phonology, such as segments, morphemes and phonological processes, are, in ET, all secondary generalizations and analogic operations, rather than primary components of representation and grammar.

Exemplar theory thus incorporates a comprehensive rejection of Neogrammarian exceptionlessness, phonemic representation, categorical and invariant grammar, and the *langue/parole* and competence/performance distinctions. It puts variation and diversity at the centre of the theory. It is inherently quantified: frequencies, proportions and contextual effects can be deduced from the distribution of variants in the exemplar cloud. But perhaps its clearest advantage for sociolinguistics is that, with modest elaboration, it can be extended to incorporate a formal representation of speakers' knowledge of the social significance of linguistic variables, including their passive knowledge of the use of others. Thus Foulkes and Docherty (2006) suggest that the memories of exemplars

also incorporate memories of who uttered them and what the social circumstances of usage were. In this way, a speaker can deduce from memory what variants are more often associated with what social context or identity. This allows them to interpret the usage of others, and to adapt their own usage accordingly. An indexical association between linguistic characteristic and social characteristic emerges statistically from the distribution of exemplars in the cloud.

Such a model addresses all the shortcomings of conventional formalisms, providing a framework for explicitly modelling inherent variability, orderly heterogeneity and social embedding. This is an attractive framework for sociolinguistics. But there are open questions about the theory concerning its capacity to explain the things that the theoretical mainstream was designed for and does well, such as categorical processes and abstract operations. Many linguistic generalizations are indeed categorical: articles precede nouns in English, without exception; many sound changes show Neogrammarian regularity: Germanic languages have no residual voiced aspirates (/bh, dh, gh/) left over from Proto-Indo-European in lexically-diffused defiance of Grimm's Law. Speakers use language in ways that go beyond the evidence, for which they have no exemplars.

For example, how does a child language learner do something new, which he or she has not encountered in the input? Consider the case of Jesse, an American-English (AmEng)-speaking child, observed by the author, who was moved to Australia at the age of 1;11. After two months of exposure to Australian English, he abruptly reformulated his pronunciation of post-tonic intervocalic obstruents: he made them all voiceless. The only exemplars in his Australian English (AusEng) input relevant to this change were AusEng pronunciations of [t] in words like *water*, *little*, which contrasted with the voiced flap in his native AmEng phonology. But rather than changing just the words for which he had such exemplars, Jesse repronounced his entire vocabulary, affecting other medial flaps, other voiced stops and even voiced fricatives: thus *daddy*>*datty*, *table*>*tapu*, *doggy*>*dockie*, *fuzzy*>*fussy*. During this period he acquired, from his AmEng-speaking parents, the words *driver* and *driving*, and duly pronounced them *drifer* and *drifing*. Aside from the intervocalic /t/ cases, he had no exemplars whatsoever for any of these new pronunciations; in fact, the input from American parents, Australian peers and childcare workers offered massive counterexamples. Jesse's behaviour suggests not a probabilistic exemplar-driven adjustment to his lexicon, but rather, an abstract reanalysis, a generalization of the form [+obs] → [-voice]/^V__V.

AQ: 'age of 1;11' means?

ET addresses such issues by recognizing processes of generalization and analogy, but these are vaguely defined and formalized. It is not clear how the generalizations across exemplars and experience that ET permits are substantively distinct from the abstract grammar it rejects. Notably, recent work in this framework takes steps to address such issues by formally recognizing abstract operations (cf. Pierrehumbert, 2006).

There are also empirical challenges to certain predictions of ET. For example, it is a central claim of the model that production is strongly influenced by such statistical properties of the input as lexical frequency. This hypothesis is in the process of being empirically tested, with mixed results to date – some studies confirm significant frequency effects on variation, while others are more equivocal. Guy et al. (2008), for example, found a frequency effect on CSD in early New Zealand English – final coronal stops in words that are more common in everyday usage were more likely to be deleted, but this result depended on using frequency counts from the corpus under study. Standard frequency counts from a large corpus of contemporary English (the CELEX corpus) did not have a significant effect.

18.7 CONCLUSION

For over a century the dominant formal theories in linguistics have paid little attention to the findings or concerns of sociolinguistics, because they treated the USE of language as fundamentally different from, and not particularly relevant to, the mental linguistic capacity of human speakers, which was the declared concern of theoretical linguistics. In place of natural language, they studied an idealization. But the limits of this strategy are now becoming clear: the idealization cannot explain the real world. If anything unifies the field of sociolinguistics, it is a shared concern with the real world, in the form of language in use. By conjoining the findings and interests of sociolinguistics with the best products of formal linguistic thinking, we can illuminate a path towards a genuine science of language.

REFERENCES

- Anttila, A. (1997) 'Deriving variation from grammar', in F. Hinskens, R. van Hout and L. Wetzels (eds), *Variation, change and phonological theory*. Amsterdam: John Benjamins. pp. 35–68.
- Anttila, A. (2002) 'Morphologically conditioned phonological alternations', *Natural Language and Linguistic Theory*, 20: 1–42.
- Anttila, A., Fong, V., Benus, S. and Nycz, J. (2008) 'Variation and opacity in Singapore English consonant clusters', *Phonology*, 25: 81–216.
- Bailey, C.-J. N. (1973) *Variation and linguistic theory*. Arlington, VA: Center for Applied Linguistics.
- Bayley, R. (1994) 'Consonant cluster reduction in Tejano English', *Language Variation and Change*, 6(3): 303–26.
- Bresnan, J., Dingare, S. and Manning, C. (2001) 'Soft constraints mirror hard constraints: voice and person in English and Lummi', in M. Butt and T. H. King (eds), *Proceedings of the LFG 01 Conference*, University of Hong Kong, online proceedings. Stanford, CSLI Publications: <http://csli-publications.stanford.edu>.
- Bybee, J. (2001) *Phonology and language use*. Cambridge: Cambridge University Press.
- Callary, R. E. (1975) 'Phonological change and the development of an urban dialect in Illinois', *Language in Society*, 4(2): 155–69.
- Cameron, R. (1993) 'Agreement, functional compensation, and nonspecific tu in the Spanish of San Juan, Puerto Rico, and Madrid, Spain', *Language Variation and Change*, 5(3): 305–34.
- Cedergren, H. J. and Sankoff, D. (1974) 'Variable rules: performance as a statistical reflection of competence', *Language*, 50(2): 333–55.
- Chomsky, N. (1964) *Current issues in linguistic theory*. The Hague: Mouton.
- Chomsky, N. (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clark, B. (2005) 'On stochastic grammar', *Language*, 81(1): 207–17.
- Clements, G. N. and Keyser, S. J. (1983) *CV phonology: a generative theory of the syllable*. Cambridge, MA: MIT Press.
- Foulkes, P. and Docherty, G. (2006) 'The social life of phonetics and phonology', *Journal of Phonetics*, 34: 409–38.
- Guy, G. R. (1980) 'Variation in the group and the individual: the case of final stop deletion', in W. Labov (ed.), *Locating language in time and space*. New York: Academic Press. pp. 251–65.
- Guy, G. R. (1991) 'Explanation in variable phonology: an exponential model of morphological constraints', *Language Variation and Change*, 3(1): 1–22.
- Guy, G. R. (1992) 'Contextual conditioning in variable lexical phonology', *Language Variation and Change*, 3(2): 223–39.
- Guy, G. R. and Boberg, C. (1997) 'Inherent variability and the obligatory contour principle', *Language Variation and Change*, 9(2): 149–64.
- Guy, G. R. and Boyd, Sally (1990) 'The development of a morphological class', *Language Variation and Change*, 2(1): 1–18.
- Guy, G. R., Hay, J. and Walker, A. (2008) 'Phonological, lexical and frequency factors in coronal stop deletion in early New Zealand English', poster presented at Laboratory Phonology 11, Wellington, New Zealand.
- Hymes, D. H. (1972) 'On communicative competence', in J. B. Pride and J. Holmes (eds), *Sociolinguistics*. Harmondsworth: Penguin.

- Kiparsky, P. (1982a) *Explanation in phonology*. Dordrecht: Foris.
- Kiparsky, P. (1982b) 'Lexical phonology and morphology', in S. Yang (ed.), *Linguistics in the morning calm*. Seoul: Hanshin. pp. 3–91.
- Kroch, A. (1989a) 'Function and grammar in the history of English: periphrastic "do"', in R. Fasold and D. Schiffrin (eds), *Language change and variation*. Amsterdam: Benjamins. pp. 133–72.
- Kroch, A. (1989b) 'Reflexes of grammar in patterns of language change', *Language Variation and Change*, 1(3): 199–244.
- Kroch, A. (2000) 'Syntactic change', in M. Baltin and C. Collins (eds), *The handbook of contemporary syntactic theory*. Malden, MA: Blackwell. pp. 629–739.
- Labov, W. (1966) *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1969) 'Contraction, deletion, and inherent variability of the English copula', *Language*, 4: 715–62.
- Labov, W. (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. and Fanshel, D. (1977) *Therapeutic discourse: psychotherapy as conversation*. New York: Academic Press.
- Labov, W. and Waletzky, J. (1967) 'Narrative analysis', in J. Helm (ed.), *Essays on the verbal and visual arts*. Seattle: University of Washington Press. pp. 12–44.
- Leben, W. (1973) 'Suprasegmental phonology'. PhD dissertation, Massachusetts Institute of Technology.
- Lim, L. T. and Guy, G. R. (2003) 'The limits of linguistic community: speech styles and variable constraint effects', *Penn Working Papers in Linguistics: Selected Papers from NWAWE 32*, 10(2): 157–70.
- Myers-Scotton, C. (1993) *Social motivation for codeswitching: evidence from Africa*. Oxford: Clarendon Press.
- Nagy, N. and Reynolds, B. (1997) 'Optimality Theory and variable word-final deletion in Faetar', *Language Variation and Change*, 9(1): 37–55.
- Nishimura, M. (1997) *Japanese–English Code-switching: syntax and pragmatics*. New York: Peter Lang.
- Paul, H. (1978 [1886]) 'On sound change', in P. Baldi and R. N. Werth (eds), *Readings in historical phonology: chapters in the theory of sound change*. Tr. by H. A. Strong from *Prinzipien der Sprachgeschichte*. University Park: the Pennsylvania State University Press. pp. 3–22.
- Pierrehumbert, J. (2001) 'Exemplar dynamics: word frequency, lenition and contrast', in J. Bybee and P. Hopper (eds), *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins. pp. 137–58.
- Pierrehumbert, J. (2006) 'The next toolkit', *Journal of Phonetics*, 34(6): 516–30.
- Poplack, S. (1980) 'Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching', *Linguistics*, 18(7, 8): 581–618.
- Sacks, H., Schegloff, E. A. and Jefferson, G. (1974) 'A simplest systematics for the organization of turn-taking for conversation', *Language*, 50: 696–735.
- Sankoff, D. (1978) 'Probability and linguistic variation', *Synthese*, 37: 217–38.
- Santa Ana, O. (1992) 'Chicano evidence for the exponential hypothesis: a variable rule pervades lexical phonology', *Language Variation and Change*, 4(3): 275–88.
- Saussure, F. de (1916) *Cours de linguistique générale*. Published by C. Bally, E. Sechehaye, and A. Reidlinger. Paris, Geneva: Payot.
- Verner, K. (1978 [1877]) 'An exception to Grimm's law', in P. Baldi and R. N. Werth (eds), *Readings in historical phonology: chapters in the theory of sound change*. Tr. by R. Stanley, from *Eine Aushnahme der ersten Lautverschiebung*. University Park: the Pennsylvania State University Press. pp. 3–22.
- Weinreich, U., Labov, W. and Herzog, M. (1968) 'Empirical foundations for a theory of language change', in W. P. Lehmann and Y. Malkiel (eds), *Directions for historical linguistics*. Austin, TX: University of Texas Press. pp. 95–195.
- Yang, C. (2000) 'Internal and external forces in language change', *Language Variation and Change*, 12(3), 231–50.