Introduction

Modern theories of linguistics rely upon the central assumption that "Language is a system." How could it be otherwise? How could we communicate with each other if there weren't rules for how we should talk and write, some sort of contract or agreement that we all share? Academic linguists are not the only ones to hold this view: in the schools, teachers of language arts prescribe for their students the rules of English (or other languages) so that they can get ahead in the world. The rules that linguists talk about are not always the same rules that language arts teachers talk about, but they both share the central assumption that there are indeed rules that help us to communicate with each other. Most people outside of universities and schools also hold the same assumption about rules. They believe that people around *here*, their neighbors, talk a certain way, and that other people from *there*, or at least not from *here*, talk a different way. American Southerners say *y'all*. Canadians say *oot* for *out*. British speakers say "the government are" when the Americans and Canadians say "the government is." That is how it is with language: the rules are a little different for people in different places or different social situations. The rules are certainly different for speakers of different languages, like Spanish and English, a lot different, so different in fact that we cannot understand each other. This fact seems to tell us that there must be rules, because speakers of different languages have such different rules that we cannot understand them. That said, even people who say that they speak our own language appear to have rules that are different enough that we just don't understand – to paraphrase the title of a famous book by Deborah Tannen (1990), which argues that men and women are like that. Still, where would be without rule systems? Men and women are different, sure, but after all, how different could they be in how they use the rules of their common language? We have the example of different languages, Spanish vs. English, or French vs. German, or any number of other different language rule systems that are not understandable to each other's speakers, to show us that without some sort of agreed-upon system we *really* would not be able to understand each other.

Yet the closer we look at how people try to understand each other, the more we find that the rules appear to be quite different from place to place, from

1

situation to situation, and from person to person. We can, for instance, recognize each other's voices, even across the room or over the telephone. The rules for a language must permit a certain amount of individual variation for this to be so. We can often recognize different social groups that people might come from, just by how they talk. For instance, we can guess how far in school someone may have gotten by how they follow the kind of rules taught in language arts classes in school. We can recognize that somebody has experience with quilting, or darts, or American football, by how they talk about the subject, by their familiarity with the vocabulary of the activity and also by their familiarity with how people usually talk about it (like "It's 180!" as the announcers say in darts matches). We can guess where people we do not know might be from, just by their use of language, whether from another country or just from another county. The rules of a language must permit variation according to regional or social groups for us to be able to guess and often be right. And we do all of these things all at the same time, from personal to social to regional evaluation of the talk we hear from the people around us. We pay attention to the *differences* in people's language as well as to the regularities of rule systems.

Furthermore, we evaluate and act upon the differences we hear, all the time. We turn our heads in a crowded room when we hear a familiar voice. We decide not to trust the advice of somebody who does not use the right language, such as a salesman who wants to sell us a television but cannot use the right words to describe what's good or bad about the choices on offer. All else being equal, we have more trust in strangers who sound like we do and may come from our home place, our kind of people, as opposed to strangers who do not sound like us. Language variation thus also plays a role in the way that we communicate with each other, along with the rules. Indeed, sometimes hearing someone break the rules is what makes the most difference in our evaluation of how a conversation is going, and so helps us to understand better how to react to it. We don't buy that new television from somebody who thinks that the "HD" in *HDTV* means 'huge-display' and not 'high-definition'!

Rule systems for languages and language variation within languages are at opposite poles of how we understand each other. Rule systems represent an ideal view, language in the abstract. Rule systems depend upon logical relationships between functional elements of language, whether those elements are features of pronunciation or lexicon or grammar. So, for example, the rules of English tell us that the word *lead* meaning 'a heavy metal' is different from the word *lead* meaning 'to guide,' even if they happen to be spelled the same, because their pronunciation is different. In terms of the pronunciation system of English, the /ɛ/ vowel (as in *bed*) is different from the /i/ vowel (as in *bee*), and we can use this systematic difference to tell words apart, and so to understand each other better. In the same way, we English speakers know that we are supposed to use pronouns like *I* or *they* when we make sentences where

pronouns are the subject (*I go downtown* or *We go downtown* or *They go downtown*); Spanish speakers usually leave out the pronoun subjects ([*yo*] *Voy al centro* or [*nosotros*] *Vamos al centro* or [*ellas*] *Van al centro*). This kind of useful logic in system making relies on the assumption, in the famous words of Noam Chomsky, of an "ideal speaker-listener, in a completely homogenous speech-community" (1965: 3). That is, we assume that all of the speakers of a language are alike in order for us to develop the system of contrasting elements, and in so doing offer an explanation of how we understand each other. Of course we are not actually all alike in practice in the way that we use our language, but that fact does not stop us from *assuming for the moment* that we are alike, or thinking that *in some way deep down* we are all alike, so that we can try to understand the rule system that helps to explain how we can understand each other. Variation in language, on the other hand, is not merely a distraction from the underlying rules; it functions and has value in its own right, as we have seen. It makes language personal, and also allows us to distinguish characteristic use of language by different groups of people. Rules and variation are opposite poles, but that does not mean that we can get rid of one. They are two sides of the same coin, and we need to consider both sides.

Our clear perception both of rules and of variation in language makes life complicated, in that it allows for quite different ways of thinking about language. Academic ideas about language consider the two poles differently, and with different emphasis. Popular ideas about language also consider the two poles with different emphasis, and differently from the way that academics think of them. Moreover, since academics who need to think about language are also brought up with the popular ideas, the popular and the academic are not entirely separate in particular cases. Conflicting ideas about language can lead to conflict in the application of ideas about language. Popular and academic ideas often come to interact on matters of public policy, such as education, and such interactions are often not as successful as one might hope because of the contrast between the popular and academic points of view.

This book proposes a model of language, called "the linguistics of speech," that attempts to provide a framework under which conflicting ideas about language might be understood for what they are. The linguistics of speech does not reject rules in favor of variation, or reject variation in favor of rules, but instead finds a place for each one in how we might think about language. By way of preview, the argument of this book for the model of the linguistics of speech includes the following ideas:

- The foundations of the linguistics of speech, as distinguished from "the linguistics of linguistic structure" that characterizes many modern academic ideas about language, are: (1) the continuum of linguistic behavior; (2) extensive (really massive) variation in all features at all times; (3) the importance of regional/social proximity to "shared" linguistic production; and (4) differential

frequency as a key factor in linguistic production both in regional/social groups and in collocations in text corpora (these points are all established with empirical study of surveys and corpora). Taken together, the basic elements of speech correspond to what has been called a "complex system" in sciences ranging from ecology and economics to physics. Order emerges from such systems by means of self-organization, but the order that arises from speech is not the same as what linguists study under the rubric of linguistic structure.

- In both texts and regional/social groups, the frequency distribution of features (variants per se or in proximate combinations, called collocations) occurs as the same curve: a "power law" or asymptotic hyperbolic curve (aka in this book, the "A-curve"). Speakers perceive what is "normal" or "different" for regional/social groups and for text types according to the A-curve: the most frequent variants are perceived as "normal," less frequent variants are perceived as "different," and since particular variants are more or less frequent among different groups of people or types of discourse, the variants come to mark identity of the groups or types by means of these perceptions. Particular variants also become more or less frequent in historical terms, which accounts for what we call "linguistic change," although of course any such "changes" are dependent on the populations or text types observed over time. In both synchronic and diachronic study the notion of "scale" (how big are the groups we observe, from local to regional/social to national) is necessary to manage our observations of frequency distributions.

- Finally, our perceptions of the whole range of "normal" variants (at any level of scale) create "observational artifacts." That is, the notion of the existence of any language or dialect is actually an "observational artifact" that comes from our perceptions of the available variants, at one point in time and for a particular group of speakers, as mediated by the A-curve. The notion "Standard," as distinct from "normal," represents institutional agreement about which variants to prefer, some less frequent than the "normal" variants for many groups of speakers, and this creates the appearance of parallel systems for "normal" and "Standard."

- The rule systems of North American academic linguistics, which we will come to call the "linguistics of linguistic structure," therefore, are related to the linguistics of speech in that language behavior, speech, is what creates the underlying distributional patterns (A-curves for all features) that yield the perceptual "observational artifacts," whether "normal" or "Standard," that we study as rule-bound systems of relationships in the linguistics of linguistic structure. Knowledge of how linguistic structure is related to language behavior, to speech, is no argument against interest in and study of structure, which will always be a useful way of looking at language. Such knowledge can help us to negotiate more effectively between different ideas in the marketplace of academic and popular notions of language.

This preview is not an argument in itself, but rather an invitation meant to help to guide readers through the course of the following chapters, in which each of its claims is developed and justified. Several central themes will emerge from this process. First, our study of language in the past has been constrained by our relative inability to store and manage evidence from speech, and modern technology (recording, computers) not only helps us to study language as we traditionally have, but it also changes the way that we can think about and model language. Second, given large bodies of stored evidence of speech, we cannot avoid quantitative methods and analysis of probabilities as a central fact about linguistic behavior. Third, our control and analysis of large bodies of speech evidence does reveal consistent principles for the organization of language behavior, at different scales of analysis, that can be assembled into an effective model for speech. The term "model" here is not the same as the traditional terms "grammar" or "language," because both of those terms assume an underlying system that makes language behavior into an object – avoidance of a priori objectification of language opens additional possibilities for how we can think about linguistic behavior. Finally, our control and analysis of speech evidence demonstrates that analysis of linguistic production alone is insufficient, and that we need also to incorporate analysis of linguistic perception in order to make an effective model of human language – and further, that linguistic perception is actually the key to the relationship between the linguistics of speech and other traditional approaches to the study of language. How we perceive language around us turns out, on the evidence of contemporary studies, to play an important role in our understanding of language itself, not just an incidental role in the evaluation of speech acts.

Unlike the point of view of some empirically oriented arguments, this book will not claim that its own approach is exclusive in order to attack competing rationalist, prescriptive, and cognitive approaches to language. An overarching theme of this book is that different approaches to language study can be well justified by the acceptance of different assumptions and priorities, and so the differing main approaches to linguistics do not really contradict each other so much as they represent different choices by their practitioners. In the market-place of ideas about language we do not now have and do not need a monopoly. We are all better off with an open market.

1 The contemporary marketplace of ideas about language

The first question that must arise for a book about "the linguistics of speech" is what we take "linguistics" to be. After we have an answer to that question, we can begin to be more specific about "the linguistics of speech." In the first chapter, we will consider contemporary ideas about language and linguistics, from both an academic and from a more popular point of view. We will see that the academic science of linguistics has not yet achieved the consensus about its basic principles that natural and physical scientists have attained for their areas of study. At the same time, the popular view of language (at least for English speakers in Britain and America) has indeed arrived at something like consensus. However, that popular view is quite different from what academic linguists think, which can lead to conflict when we need to make decisions about language and public policy, as in educational policy. This contemporary competition of ideas about language can be described as a marketplace, in which ideas about language are promoted and accepted, bought and sold. In order to understand "linguistics," and thus to prepare the way for a discussion of "the linguistics of speech," we need to try to understand what motivates the buyers and the sellers in the marketplace. For our purposes this will not mean a minute examination of academic theories or popular beliefs about language, but instead a sketch of the main differences between ideas so that we can observe the interaction of the ideas in a test case, the Ebonics controversy, in which the conflict of ideas becomes most clear.

The academic marketplace of ideas about language

Two different basic approaches to the creation of rule systems have been popular in modern American academic linguistics, and these approaches apply the assumption of a homogenous speech community in different ways. Under the first approach, *structuralism*, linguists gather information about a language from one or two or some small number of speakers, and attempt to describe the system of the language from what they say. It is not necessary to talk to more than a few speakers, perhaps just one, because the structuralist assumes that the speakers of a language are more or less alike in that they share a

6

rule system. Such was the case at the beginning of American linguistics, when linguists like Leonard Bloomfield described the system of Native American languages on the basis of conversation with just a few speakers. The same was true of the Army Language Program, on which a great many American linguists collaborated during World War II in order to prepare dictionaries and grammars of the many languages of the world that English-speaking servicemen would encounter: Raven I. McDavid, Jr., otherwise best known for his work on American English, told me (p.c.) about writing the Army materials on Burmese on the basis of a single speaker seated in a chair by his desk in New Haven. In practice, structuralist descriptions of a language tend to get larger and more complex as linguists talk to more people and hear more and different details about the language. So, for example, *A Comprehensive Grammar of the English Language*, the famous "London School" grammar of English by Randolph Quirk, Sidney Greenbaum, Geoffry Leech, and Jan Svartvik (1985), requires 1779 large-format pages, nearly five pounds of book, to describe the grammar of English. Even so, ever since its publication linguists have been writing scholarly articles about facets of English that the book does not describe. The size of the *Comprehensive Grammar* pales beside the magnificent scale of the multi-volume *Oxford English Dictionary*, which does not even claim to list and define all the words of English (it leaves out proper names and chemical names, for instance). There is no complete description anywhere of every aspect of English grammar or of every word in the English vocabulary. No single person actually uses all of the grammar in the *Comprehensive*, or all of the words and senses in the *OED*, but that is not the point of a structuralist description of English. The assumption of an essentially homogenous speech community allows structuralists to create a description of the rule system of a language on the basis of a small number of speakers, and to enhance and improve the description as more evidence is collected, because the system of a language is independent of any particular individual speaker. To "share" the system does not mean that individuals have to embody or control every aspect of it, any more than individual bank customers own all the money in the banking system that they share with other customers. Individuals take part in the banking system with their own money, and we can get a fair idea of the system from one person's transactions; we can get to know more about the banking system per se as we look at more people's transactions. Structural linguistics improves its description of the system of a language as it integrates evidence from more speakers.

The other major modern American approach to linguistics is *generativism*. Generativists take the creation of rule systems, by means of the assumption of a homogenous speech community, to address our human capacity for speech. Generativist study of rule systems contributes to the description of a "universal" system upon which we draw in the formation of the rule system for our own

particular language. Chomsky and generativism swept into American linguistics with the "linguistic shot heard 'round the world" (another comment from Raven McDavid (p.c.)), Chomsky's 1957 book *Syntactic Structures*. By that time American structuralism had entered a difficult period. The complexity of structuralist generalizations had to increase as more and more evidence for English and other languages piled up and required integration with existing structural descriptions of language systems. This often led to vigorous competition between linguists who offered different descriptions for the complex facts. Generativism, on the other hand, focused on the creation of the smallest possible rule systems that could "generate" the acceptable sentences of a language, according to grammaticality judgments of its speakers. Chomsky proposed an "evaluation metric" by which competing grammars could be justified, and explained that an evaluation metric is necessary if linguists seek to develop not simply a description of a particular language, but to extrapolate from particular languages towards a theory of natural language in general (1965: 41; for his early programmatic treatment of rules and grammars, see Chomsky 1961). Chomsky suggests that the central problem is deciding which generalizations are significant (1965: 41):

We have a generalization when a set of rules about distinct items can be replaced by a single rule (or more generally, partially identical rules) about the whole set, or when it can be shown that a "natural class" of items undergoes a certain process or set of similar processes. Thus, choice of an evaluation measure constitutes a decision as to what are "similar processes" and "natural classes" – in short, what are significant generalizations.

Chomsky called generalizations that render the description of a system more complex "spurious generalizations" as opposed to significant ones, which of course reversed the structuralist process of evaluation that always led to greater complexity. The smaller rule system is by definition better in generativism, which in time drove generativist theory towards "minimalism" (as in Chomsky's 1995 *The Minimalist Program*). Structuralist descriptions started small and got bigger, while those of generativism started bigger and got smaller, all the while under the assumption that there must be a system of rules to help explain how the speakers of a language understand each other.[1]

Language variation, on the other hand, arises from language in use, from how people actually speak and write. The study of language in use for itself, not for the purpose of discovering any abstract rule system, does not assume the existence of a homogenous speech community – indeed it assumes that everybody and every group and every place, every situation, is different. The goal of

---

[1] Kretzschmar and Celis (1998) proposed that a better model of language would entertain "mid-level generalizations" as opposed to minimalist or global structural rules – the largest or smallest possible grammars are not the only possibilities. Structuralism and generativism are not the only choices in the marketplace, just the best known of many choices.

such study is not to reveal an underlying system that speakers share, but instead to characterize what speakers actually do with language. *Empirical linguistics* is one name for the study of language variation, although it is certainly true that both structuralism and generativism have an empirical component. Victoria Fromkin, late author of a best-selling textbook with a generativist approach (most recent edition, 2008), once complained to me (p.c.) about this use of "empirical linguistics" to describe such study of language variation, pointing out quite rightly that generativists must have some "empirical" contact with language in order to pursue their own approach. Structuralism, too, relies on empirical findings, at first from few speakers but eventually integrating evidence from many more, in order to pursue structural generalizations about a language. Nonetheless, the term *empirical linguistics* has emerged as a description for the kind of modern linguistics that begins with language in use and not with any assumption of rule systems. Geoffrey Sampson writes in a book actually called *Empirical Linguistics* that (2002: 1)

Language is people talking and writing. It is a concrete, tangible aspect of human behaviour. So, if we want to deepen our understanding of language, our best way forward is to apply the same empirical techniques which have deepened our understanding of other observable aspects of the universe during the four centuries since Galileo.

Yet Sampson finds that "[his research] suggests that one should not talk about different grammars for fiction or technical writing. Instead we need to think in terms of a single grammar, which generates a range of tree structures, some large and some small" (2002: 35). Sampson, though a British scholar, might best be categorized as one interested in the North American practice of Natural Language Processing (NLP), which features computational approaches to large corpora that focus on rule-based regularities, and whose object is the generation of tree banks of grammatical structures. Sampson creates an aggressively argued opposition between the "intuition" that he attributes to Chomsky and generativism, and the scientific method of hypothesis development and testing that he wants linguistics to share with the physical sciences. One might argue in reply that American structuralism was actually a good example of hypothesis development and testing, and that its practitioners nonetheless created the conditions under which Chomsky's approach could flourish by their apparent failure to find a way forward through the necessity for increasingly complex generalizations demanded by the weight of the evidence. Chomsky himself claims modern scientific procedure for generativism, and says that the preference for large collections of language data "is just a misunderstanding of the notion of empirical" (Andor 2004: 97–98). It does little good to argue about what "empirical" or "empiricism" might mean, when in fact both sides claim to be engaged in the scientific inquiry of hypothesis development and testing, even one with empirical contact with the language, just with a different means for

evaluation of evidence (for more on this debate, see Kretzschmar 2006a). What is really different from structuralism and generativism about the study of language variation, of language as it is actually spoken by real people, comes not from paying empirical attention to what people say, but from the absence in language variation study (what will in this volume come to be called "the linguistics of speech") of any assumption of rule systems or of ideal speaker-listeners in completely homogenous speech communities.

Within North American linguistics, the area of sociolinguistics seems to be the most likely branch to concern itself with language variation. However, Labovian sociolinguistics is firmly associated with rule systems, in particular with "the vernacular." Shana Poplack offers a concise description (1993: 258):

The primary object of description of the [sociolinguist] is the speech of individuals qua members of a speech community, i.e. informants specifically chosen (through ethno-graphic or sociological methods) to represent the major axes of community structure … A specific goal of this procedure is to gain access to the vernacular, the relatively homogeneous, spontaneous speech reserved for intimate or casual situations. This is taken to reflect the most systematic form of the language acquired by the speaker, prior to any subsequent efforts at (hyper-) correction or style-shifting (themselves imposed by the combined pressures of group membership and the social meaning within that group of the linguistic options available).

Poplack's terms are clearly associated with language as system. As she reports, sociolinguists take an interest in what people say not for their personal language behavior, but as individuals who may be chosen to represent collectivities (speech communities) that are assumed to exist. Sociolinguists are interested in "the most systematic form of the language," before it can be deflected by the messy details of human social organization. Labov is reported to have resisted the term "sociolinguistics" for his own work, preferring to think that he simply was doing "linguistics" (Trudgill 1984: 2–3). Labov's monumental volumes on *Principles of Linguistic Change* (1994, 2001) are defenses of the language-as-system approach to large collections of language evidence, historical evidence in the first volume, especially from New York, and contemporary evidence, especially from Philadelphia, in the second (see further Kretzschmar 1996b, 2005). Once more, as for Sampson, the empirical collection of speech evidence is not the same as analysis of language variation on its own merits, since it is perfectly possible, though controversial in contrast to those who provide isolated examples of usage, to collect bodies of speech evidence and to analyze the evidence in terms of rule systems.

While the structuralists and generativists pursued their approaches to linguistics in America, Firth, Halliday, Sinclair, and others were taking a quite different approach in Britain. While American linguists began with the idea of structured rule systems, British linguists began with conversation: