Chapter 1

Overview

The investigation of language can be carried out in many different ways. For example, we might learn to speak a language as a way to understand it. We might also study literature and poetry to the same end. We might even write novels and poetry to understand the intricacies of expression that some language provides.

These are valuable pursuits, not to be denigrated in any way, but they do not provide for a scientific understanding of language, one where we can make falsifiable or testable claims about our object of study. A falsifiable claim is one that can be disproven with real data. For example, if we foolishly hypothesize that Shakespeare wrote good poetry because he wrote in English, we would need some objective way to assess how “good” some piece of poetry is independent of the language the author wrote in. If the hypothesis rested instead entirely on our own ideas about how good individual poems are, then it would surely not be falsifiable, and thus not be science.

1.1 A Scientific Approach

There are, of course, a number of ways to do science with language. We might investigate the range of sentence types that some speaker can produce or that some corpus contains. For example, do male or female characters in Shakespeare have longer sentences? We might look at the set of sounds that can occur in any one language, or the set of sounds used in some specific poetic context. Do the languages of Europe have more vowels than the languages of India? We might want to investigate the time course of language
acquisition or of language change. Do children learn all the vowels of English before or after they learn all the consonants? Are vowels or consonants more “stable” over the history of the English language?

As we lay out these different questions, we beg the question of why we might expect these different outcomes. Why do we believe that language should work in any particular way? Our expectations about how language should behave are our theory of language. For example, if we hypothesize, for example, that all languages have the vowel [a], as in English father, then we have an implicit theory about how language works and about the vowel [a]. If we believe that male or female characters have longer sentences in Shakespeare’s plays, then we have a theory of sentence length and its relationship to gender.

These expectations can be fairly informal and intuitive. For example, we might believe that [a] is a very easy vowel to produce and that languages make use of easier sounds before making us of harder sounds. With respect to sentence length, we might have some idea about the relatively minor role female characters played in Shakespeare and how that would be reflected in the language of those characters.

As we proceed along these lines, as we make, test, and refine our hypotheses about language, it behoves us to make our theory of language more explicit. The more explicit our theory of language is, the more falsifiable it is. As scientists, we want our theory of language to be as testable as possible.

As our theory becomes more explicit, it tends to become more formal. A formal theory is one where the components of the theory are cast in a restricted metalanguage, e.g. math or logic. The languages of math and logic are quite precise. A theory cast in those terms can make very specific predictions and those be as falsifiable as possible. A theory cast in normal English can be quite vague. An analogy would be the language of contracts. Typically, such documents are quite hard to make out for laymen as they are written using very specific language that lawyers have very specific interpretations for. While this kind of language can be quite frustrating for the rest of us, it is essential. The legal metalanguage of contracts provides a mechanism whereby our rights and commitments can be negotiated in contracts.

Likewise, mathematical and logical formalisms can appear daunting as parts of a theory of language, but they can be indispensable to making a theory maximally falsifiable.

Understanding the formalisms used in theories of language requires specialized knowledge of various formal domains. That is the point of this book.
1.2 Other Reasons for Formal Theories

Intuitively, formalism is all the symbols, all the stuff that looks like math. As we've discussed in the previous section, the formalism is just a mechanism for being maximally explicit. If we want to build theories that we can test in precise ways, then we need to have theories that are as precise as possible. A proper formalization enables us to do this.

Formalization also enables us to implement our theories computationally. We might want to write a computer program that mimics some aspect of the grammar of a language, its sentences, words, or sounds. For example, imagine we wanted to write a program that would produce poetry. We would supply a vocabulary and the program would spit back a random poem. This may seem rather silly, but is, in fact, a hugely complicated task. We would have to provide the program with some definition of what constitutes poetry. To the extent that this definition had any content, the task becomes quite complex. For example, how do we define “rhyme”? How do we get the program to produce grammatical sentences and not just random strings of words? How do we get the program to figure out how words are actually pronounced? Computer programs are merciless in requiring specific answers to questions like these and an inexplicit theory of language will be of little help.

Why would we want to write such programs? There are two broad reasons. The first is that it is another way to test our understanding of the theory. For example, if we have the wrong theory of rhyme, then our program would produce bad poetry. The second is that we may actually want to do something useful in the outside world with our theories, use them for some other purpose other than the pursuit of “truth”. While a program that wrote poetry would seem of little use, a program that examined text and identified it as poetry or not might be of great use. In either case, being as explicit as possible in our formalization of our theory makes implementing that theory as painless as possible.

Another reason why we formalize theories of language is to understand the general character of formalization better. For example, if we challenge ourselves to characterize some aspect of language in terms of first-order logic\(^1\), we may find out something about logic too. Thus formalization of theories of language can also tell us about math and logic more generally.

\(^1\)More on this in chapters 4 and 5 below.
1.3 An example

Let’s consider a simple example. Pretty much anyone would agree that “two times two equals four” is the same thing as $2 \times 2 = 4$. In a relatively simple domain like multiplication, there does not appear to be much to be gained by translating words into an equation. Consider on the other hand a sentence like the following:

(1.1) Two women saw two men.

If we are interested in the meanings of sentences, then a bit of formalism might be helpful in characterizing what such a sentence can mean. In the case at hand, this sentence has at least two different interpretations. First, the sentence could mean that there are precisely two women and precisely two men and each of the two men were seen by at least one of the women. However, there is another possible interpretation where there are only two women, but up to four men. Each woman saw two men, but they may not have been the same men. We can make this a little clearer if we assign names to the individuals: Mary, Molly, Mark, Mike, Mitch, and Matt.

(1.2) First interpretation: Mary and Molly saw Mark and Mike.

(1.3) Second interpretation: Mary saw Mark and Mike, and Molly saw Mitch and Matt.

The second interpretation allows for other possibilities as well. Thus characterizing the meaning of a sentence might require some fairly elaborate formal machinery.

1.4 Organization of the Book

The main topics we cover in this book are the following.

Set Theory Abstract theory of elements and groups. Underlines virtually everything else we’ll treat.

Logic A formal system with which we can argue and reason. There are two ways to look at this. One way to look at it is as a formal system
that underlies how we make arguments. However, another way to look at it is as a “perfect language” with syntax and semantics rigorously defined.

**Formal Language Theory** An explicit way to look at languages and the complexity of their description with an eye to how we might compute things about those languages.

**Probability** How to be explicit about degrees of likelihood, essential for experimental disciplines, linguistic behavior, language change, sociolinguistics, . . . .

1.5 Acknowledgements

Thanks to Andrew Carnie and Diane Ohala for useful discussion. All errors are, of course, my own.