Part I

Introduction

1

Linguistics:
The Scientific Study of Human Language

1.0 Introduction

Human language, that unique characteristic of our species, has been of interest throughout history. The scientific study of human language is called **linguistics**. A **linguist**, then, is not someone who speaks many languages (although many linguists do); such individuals are **polyglots**. A linguist is a scientist who investigates human language in all its facets, its structure, its use, its history, its place in society.

The form and structure of the kinds of linguistic knowledge speakers possess is the concern of theoretical linguistics. This theory of **grammar** – the mental representation of linguistic knowledge – is what this textbook is about. But the field of linguistics is not limited to grammatical theory; it includes a large number of subfields, which is true of most sciences concerned with phenomena as complex as human language.

A bird's-eye view of the field

Theoretical Linguistics (the concern of this textbook), often referred to as **generative linguistics,** has its basis in views first put forth by Chomsky's 1955 *The Logical Structure of Linguistic Theory.* In this and the subsequent books and articles by Chomsky and those that embraced these views, a major aim was to characterize the nature of human linguistic knowledge or **competence** (represented in the mind as a mental **grammar**); that is, to explain or account for what speakers know which permits them to speak and comprehend speech or sign (the languages of the deaf). The production and comprehension of speech is referred to as **performance,** distinct from competence but dependent on it.

Descriptive linguistics provides analyses of the grammars of languages such as Choctaw, Arabic, Zulu. 'Indo-European-linguistics,' 'Romance linguistics,' 'African linguistics,' refer to the studies of particular languages and language families, from both historical and synchronic points of view.

Historical linguistics is concerned with a theory of language change – why and how languages develop. The **comparative method,** developed in the nineteenth century by such philologists as the brothers Grimm and Hermann Paul, is a method used to compare languages in the attempt to determine which languages are related and to establish families of languages and their roots.

Anthropological or **ethno-linguistics** and **sociolinguistics** focus on languages as part of culture and society, including language and culture, social class, ethnicity, and gender.

Dialectology investigates how these factors fragment one language into many. In addition, **sociolinguistics** and **applied linguistics** are interested in language planning, literacy, bilingualism, and second language acquisition. Applied linguistics also covers such areas as discourse and conversational analysis, language assessment, language pedagogy.

Computational linguistics is concerned with natural language computer applications, e.g. automatic parsing, machine processing and understanding, computer simulation of grammatical models for the generation and parsing of sentences. If viewed as a branch of Artificial Intelligence (AI), computational linguistics has the goal of modeling human language as a cognitive system.

Mathematical linguistics studies the formal and mathematical properties of language.

Pragmatics studies language in context and the influence of situation on meaning.

Neurolinguistics is concerned with the biological basis of language acquisition and development and the brain/mind/language interface. It brings linguistic theory to bear on research on aphasia (language disorders following brain injury) and research involving the latest technologies in the study of brain imaging and processing (CT, PET, fMRI, MEG, ERP).

Psycholinguistics is the branch of linguistics concerned with linguistic performance – the production and comprehension of speech (or sign). An area of

psycholinguistics, which in some ways is a field in its own, is child **language acquisition** – how children acquire the complex grammar which underlies language use. This is a subject of major concern, particularly because of the interest in the biology of language. (This topic will be covered in the text because of its relevance to theories of grammar.)

There are textbooks which deal with each of these subfields, among others. Basic to all of them is the nature of language itself, the subject of this textbook.

1.1   Panini to Chomsky and After

The interest in the nature of human language appears to have arisen when the human species evolved in the history of time. There is no culture that has left records that do not reveal either philosophical or practical concerns for this unique human characteristic. Different historical periods reveal different emphases and different goals although both interests have existed in parallel.

Egyptian surgeons were concerned with clinical questions; an Egyptian papyrus, dated ca. 1700 BCE, includes medical descriptions of language disorders following brain injury. The philosophers of ancient Greece, on the other hand, argued and debated questions dealing with the origin and the nature of language. Plato, writing between 427 and 348 BCE, devoted his *Cratylus Dialogue* to linguistic issues of his day and Aristotle was concerned with language from both rhetorical and philosophical points of view.

The Greeks and the Romans also wrote grammars, and discussed the sounds of language and the structures of words and sentences. This interest continued through the medieval period and the renaissance in an unbroken thread to the present period.

Linguistic scholarship, however, was not confined to Europe; in India the Sanskrit language was the subject of detailed analysis as early as the twelfth century BCE. Panini's Sanskrit grammar dated ca. 500 BCE is still considered to be one of the greatest scholarly linguistic achievements. In addition, Chinese and Arabic scholars have all contributed to our understanding of human language.

The major efforts of the linguists of the nineteenth century were devoted to historical and comparative studies. Ferdinand de Saussure (1857–1913), a Swiss linguist in this tradition, turned his attention instead to the structural principles of language rather than to the ways in which languages change and develop, and in so doing, became a major influence on twentieth-century linguistics.

In Europe and America, linguists turned to descriptive synchronic studies of languages and to the development of empirical methods for their analysis. Scholars from different disciplines and with different interests turned their attention to the many aspects of language and language use. American linguists in the first half of the century included the anthropologist Edward **Sapir** (1884–1939), interested in the languages of the Americas, language and culture, and language in society, and Leonard **Bloomfield** (1887–1949), himself an historical and comparative linguist, as well as a major descriptive linguist who emerged as the most influential linguist in this period. Both Sapir and Bloomfield were also concerned with developing a general theory of language. Sapir was a 'mentalist' in that he believed that any viable linguistic theory must account for the mental representation of linguistic knowledge, its 'psychological reality'; Bloomfield in his later years was a follower of behaviorism, which was the mainstream of psychological thought at the time, a view that precluded any concern for mental representation of language and, in fact, for the mind itself.

In Europe, Roman Jakobson (1896–1982), one of the founders of the Prague School of Linguistics, came to America in 1941 and contributed substantially to new developments in the field. His collaboration with Morris Halle and Gunnar Fant led to a theory of **Distinctive Features** in phonology, and Halle has remained one of the leading phonologists of the last decades. In England, phoneticians like Daniel Jones (1881–1967) and Henry Sweet (1845–1912) (the prototype for G. B. Shaw's Henry Higgins) have had a lasting influence on the study of the sound systems of language.

In 1957 with the publication of *Syntactic Structures*, Noam Chomsky ushered in the era of generative grammar, a theory which has been referred to as creating a scientific revolution. This theory of grammar has developed in depth and breadth. It is concerned with the biological basis for the acquisition, representation and use of human language and the universal principles which constrain the class of all languages. It seeks to construct a scientific theory that is explicit and explanatory.

The chapters that follow are based to a great extent on the developments in linguistic theory that have occurred since the publication of *Syntactic Structures* in 1957 and *Aspects of the Theory of Syntax* in 1965. In subsequent years, Chomsky has continued to develop his theory in such major works as *Remarks on Nominalization* (1970), *Conditions on Transformations* (1973), *Lectures on Government and Binding* (1981), *Barriers* (1986), *Principles and Parameters in Syntactic Theory* (1981), and *The Minimalist Program* (1995).

In the following chapters, basic notions in these publications as well as many others in areas other than syntax are presented at an introductory and basic level.

1.2   Aims of Linguistic Theory

Three key questions were posed by Chomsky in 1986 which remain pivotal in linguistics today:

What constitutes knowledge of language? (Competence)
How is knowledge of language acquired? (Acquisition)
**How is knowledge of language put to use? (Performance/language
   processing)**

As stated above, this text will be primarily concerned with the first question viewed in relation to the second. The development of language from infancy provides insights into the nature and structure of language itself and therefore is discussed in each part. An understanding of language use (performance), the main tenet of psycholinguistic research, depends on our understanding of what is being put to use. We will discuss the distinction between linguistic knowledge (**competence**) and use (**performance**) below.

1.3   What Constitutes Knowledge of Language? Grammar as the Representation of Linguistic Competence

Knowledge of a language permits one to connect sounds (or gestures in sign languages) with meanings, that is, to understand a spoken or signed utterance, and to express our thoughts through speech or signs. Note that the **sign languages** of the deaf are basically the same as spoken languages, using a gestural/visual modality instead of the sound/aural perceptual modality of speech. Except where specifically referring to speech sounds, discussion of the nature and characteristics of language should be interpreted as referring to both spoken and signed languages.

Linguistic knowledge as represented in the speaker's mind is called a **grammar**. Linguistic theory is concerned with revealing the nature of the mental grammar which represents speakers' knowledge of their language.

If one defines grammar as the mental representation of one's linguistic knowledge, then a general theory of language is a theory of grammar. A grammar includes everything one knows about the structure of one's language – its **lexicon** (the words or vocabulary in the mental dictionary), its **morphology** (the structure of words), its **syntax** (the structure of phrases and sentences and the constraints on well-formedness of sentences), its **semantics** (the meaning of words and sentences) and its **phonetics** and **phonology** (the sounds and the sound system or patterns). A theory of

grammar specifies the nature of each of these components and the universal aspects of all grammars.

Each of these different kinds of linguistic knowledge constitutes a component of the mental grammar. But what kind of knowledge is this? What do speakers know? First it must be noted that we are not speaking of conscious knowledge. Most of us (before taking a course in linguistics) are totally unaware of the extent of our tacit unconscious knowledge of our language. We have no idea of the complexity of this knowledge. Some of this complex knowledge will be revealed in the chapters to come. As a way of introduction, however, we can illustrate the nature of this linguistic competence as represented in the components of our mental grammars.

1.3.1   The lexicon

Every speaker of language has a dictionary or **lexicon** in their head, with all the words which they know, words like *cat, witch, cauldron, Macbeth, jester, vocabulary, slay, betray, love, hate*. It has been estimated that the average person knows from 45,000 to 60,000 words; these must be stored in the mental lexicon (see chapter 2).

1.3.2   Morphology

A speaker of a language also knows how words are structured, that, for example, in English, words are composed of bare **roots** like *witch*, or roots with **suffixes** like *witch-es* or words with **prefixes** like *dis-en-chant* or words with prefixes and suffixes like *dis-en-chant-ment*, and furthermore they know that these parts of words must occur in a certain order. That is, to put a suffix like *ment* at the beginning of a word – **mentenchant* – would make it **unacceptable** or **ungrammatical**. (Throughout the book, linguistically ill-formed – unacceptable, or ungrammatical – words or phrases or sentences will be preceded by an **asterisk ***.) Since speakers can distinguish between acceptable and unacceptable forms, that is, they accept *lover* but reject **erlove*, for example, then this is part of our grammatical knowledge, represented in our mental grammars.

1.3.3   Syntax

Part of our linguistic knowledge tells us what constitutes a well-formed string of words, how to put words together to form phrases and sentences. We know when such strings of words are grammatical (**well-formed**) or ungrammatical (**ill-formed**), as in the difference between (1) and (2):

(1)   Lear had three daughters.

(2)   *Had three Lear daughters.

Note that grammaticality does not depend on our having heard the sentence before or whether it is true, since (3)

(3)   Lear had only two daughters.

is a grammatical sentence but according to the Shakespeare tragedy it is not a true sentence. And since Lear is a character in a play and does not nor did not exist as a King with three daughters, the acceptability of (1) does not depend on whether there is a referent in the real world for the information being conveyed.

Our knowledge of syntax (and semantics) also accounts for the fact that we know that (4) is an ambiguous sentence (a sentence with more than one meaning).

(4)   Cordelia loved Lear more than Regan.

Although one meaning may come to mind first upon our hearing such a sentence, perhaps

(4)   a.   Cordelia loved Lear more than Regan loved Lear.

speakers of English also know that (4b) may be the meaning of (4):

(4)   b.   Cordelia loved Lear more than Cordelia loved Regan.

The nature of syntax also accounts for the fact that there is an unlimited – infinite – set of sentences in any language. One cannot put a limit on the length of a sentence and thus cannot put a limit on the number of sentences.

(5)   Lear loved Cordelia.

(6)   Lear loved Cordelia and Cordelia loved Lear.

(7)   Lear, who had three daughters, loved Cordelia the most.

(8)   Lear, who loved Cordelia, became very angry with her when she would not tell him in words how much she loved him and that made him cut her off without any lands or riches, which pleased Regan and Goneril but was very unfair because Cordelia really loved him the most and was unselfish and kind and all of this led to the terrible tragedy of King Lear.

We are able to embed sentences within sentences as shown in (9a–c):

(9) a. Cordelia was Lear's youngest daughter.
    b. She loved him.
    c. Cordelia, who was Lear's youngest daughter, loved him.

We are also able to conjoin sentences as in (10):

(10) Cordelia was Lear's youngest daughter and she loved him.

We know how to negate sentences and form questions.

(11) Cordelia was not Lear's youngest daughter.

(12) Was Cordelia Lear's youngest daughter?

We also know that in (13) *him* cannot refer to Lear, but in (14) *him* can refer to Lear or to someone else:

(13) Lear loved him.

(14) Regan told Lear she loved him.

   Chapters 3, 4, and 5 will deal with some of these questions and many more that have to do with our tacit knowledge of syntactic structure, knowledge which also must be represented in the mental grammar to account for our ability to make such judgments.

1.3.4 Semantics

Speakers also know quite a lot about what the expressions in their language mean or signify, and it is this knowledge which makes the patterns of sounds or gestures 'symbolic'. A sentence is like a sort of acoustic or gestural picture – it represents something – though the way they manage to be representational is different from pictures. For example, we know that the spoken or written word *Shakespeare* can be used to name a person; and we know that the spoken or written phrase *wrote plays* signifies a property that some people but not others have, and we can put these two ideas together to recognize the meaning of a sentence like:

(15) Shakespeare wrote plays.

The simple idea that the subject names something and the rest of the sentence tells us about that thing is not quite right, though. For example, this does not work for sentences like

(16)   No witches wrote plays.

since sentence (16) is perfectly meaningful even though *no witches* does not refer to any particular person. In chapters 7, 8, and 9 we will provide an account of the meanings of sentences which handles both (15) and (16) equally well.

Furthermore, we can understand sentences that are built up in all the ways that the syntax allows: we understand sentences that are embedded and coordinated; we understand negated sentences and questions. Our ability to recognize syntactic structures comes with this parallel ability to understand what the infinite range of structures signifies.

When we consider the meanings of expressions, we also notice that similar meanings can sometimes be conveyed in very different ways. For example, the following two sentences have very similar meanings:

(17)   Usually, a witch does not write plays.

(18)   Most witches do not write plays.

This kind of knowledge represents semantic knowledge in our mental grammars and will be discussed in chapters 7, 8, and 9.

1.3.5   Phonetics and phonology

Speakers' knowledge of their language also includes knowledge of the sounds and sound patterns which occur. We know what sounds are in the language and what sounds are not. Speakers of English know, unconsciously for the most part, that there are more than five vowel SOUNDS in the language, as shown by the vowel sounds which differentiate the following words from each other: *bit, beat, bet, bait, bat, boot, but, boat, bought, put, pot*. We use five LETTERS – *a, e, i, o, u* – to represent these different vowel sounds in our writing system. We see that there is no one-to-one mapping between alphabetic symbols and the sounds they represent.

Speakers of English also know what strings of sounds are words, are possible words, and are impossible words in their language. Thus, *clasp* occurs in most speakers' mental dictionaries, while *clisp* or *klisp* does not (since no meaning is 'attached' to these sounds), but this nonsense form could become a word since it does not violate any constraints on sequences of sounds that are permitted in English. But **lkisp* is not a possible word in English nor is **ngisp* since in the first case, the sequence of the sounds *lk* cannot begin a word, and in the second case, *ng*, the sound which ends the word *king*, cannot begin a word. These are not 'rules' or 'laws' established by writers of textbooks but are constraints on the sound patterns of

language which children learn when they acquire their language. Such constraints and additional information about speech sounds and sound patterns, discussed in chapters 11 through 14, are part of our phonological knowledge represented in the mental grammar of English.

1.4 Mental Grammar, Universal Grammar, Descriptive Grammars, Teaching Grammars, and Prescriptive Grammars

Grammar as viewed here is different from the usual notion of grammar. When viewed as the representation of a speaker's linguistic competence, a grammar is a mental system, a cognitive part of the brain/mind, which, if it is one's first native language, is acquired as a child without any specific instruction. The word *grammar* is often used solely in reference to syntax. But we use it to refer to all aspects of linguistic competence. In addition to its use as referring to the mental system, when linguists describe this knowledge shared by a language community, the description is also called the grammar of the language. Of course no two speakers of a language have identical grammars; some may know words that others do not, some may have some idiosyncratic rules or pronunciations. But since they can speak to each other and understand each other there is a shared body of knowledge, which is what we are calling their mental grammars. **Descriptive grammars** are thus idealized forms of the mental grammars of all the speakers of a language community.

  The grammars of all languages are constrained by universal 'laws' or 'principles,' a view which differs from that of many linguists in the pre-Chomsky period some of whom held that languages could differ in innumerable ways. The more we look at the languages of the world, the more support there is for the position taken by Roger Bacon, a thirteenth-century philosopher, who wrote:

> He that understands grammar in one language, understands it in another as far as the essential properties of grammar are concerned. The fact that he can't speak, nor comprehend, another language is due to the diversity of words and their various forms, but these are the accidental properties of grammar.

There is much evidence to support this view, which today is based on the recognition that there is a biological basis for the human ability to acquire language. The child enters the world with an innate predisposition to acquire languages which adhere to these universal principles, that is, a

genetically determined mental system which is referred to as **Universal Grammar** or **UG**. This will be discussed further below.

While UG constrains the form of the grammars of all human languages, there are the 'accidental differences' which constitute cross-linguistic variability. Using the theory of grammar which specifies UG, linguists investigate and analyze specific languages and construct the descriptive grammars, mentioned above. Thus, while UG may specify that the sound represented by *th* beginning the word *thane* (as in 'Macbeth, the *thane* of Cawdor') or *thigh* is a possible speech sound, the descriptive grammar of English will include this sound in its grammar but the grammar of French will not. Some languages permit the dropping of a pronominal subject, others do not. In some languages one complex word may be equivalent to a whole sentence in another language. These differences will be revealed in the descriptive grammars of these languages. Descriptive grammars also may serve as the basis for **teaching** or **pedagogical** grammars which are used to teach someone a second language or a variation (dialect) of one's native language.

Descriptive grammars aim at revealing the mental grammar which represents the knowledge a speaker of the language has. They do not attempt to prescribe what speakers' grammars should be. While certain forms (or dialects) of a language may be preferred for social or political or economic reasons, no specific dialect is linguistically superior to any other. The science of linguistics therefore has little interest in **prescriptive grammars**.

It should also be noted, as discussed in chapter 2, that the majority of languages of the world have no written form. They are, however, as complex and rational as languages with a written **orthography** or alphabet. Speech (or sign) is primary, part of the natural endowment of the human species; writing systems are derived from spoken languages, which is why every normal human anywhere in the world who receives linguistic input as a child will learn the language of the environment but will not necessarily learn to read or write unless being specifically taught. Even deaf children deprived of linguistic input 'invent' their own gestural language, which develops and changes to conform to the constraints of UG. But such children do not construct written languages on their own.

1.5 How is Knowledge of Language Acquired? The Logical Problem of Child Language Acquisition

Young children, limited in so many respects, accomplish with apparent ease a remarkable cognitive feat. In just a few short years, without benefit of any direct instruction or correction, they develop a very complex and uniform cognitive system of linguistic knowledge, a grammar of the

language being acquired. Just how children do this is a central question that linguistic theory tries to answer. What makes the acquisition problem particularly intriguing is that we come to know vastly more about our language than we are provided evidence for in our linguistic environment. Consider, for example, the sentences in (19) and (20):

(19)   Polonius is eager to please.

(20)   Polonius is easy to please.

Superficially, these sentences are very similar. Only the adjective is different, *easy* in (20) and *eager* in (19). Yet, upon reflection an English speaker will know that they differ in a way that goes beyond the choice of adjective. In (19) the subject *Polonius* is the pleaser, while in (20) Polonius is the one being pleased and someone else, unspecified in the sentence, is the pleaser. The sentence in (20) might be paraphrased as 'It is easy to please Polonius,' while the sentence in (19) has no such paraphrase. *It is eager to please John* is not a good sentence in English. What we as speakers know about such sentences goes beyond what is exemplified in the superficial form.

   Similarly, the sentence in (21) has two meanings associated with it:

(21)   Visiting witches can be dangerous.

Speakers of English know that this sentence can mean (a) that it is dangerous to visit witches, or (b) witches who are visiting are dangerous. The ambiguity of such a sentence is not given in the superficial form of the sentence. Yet, it is something we know. Just as we know that the sentence in (22) is a grammatical question in English while the sentence in (23) is not, though (22) and (23) mean essentially the same thing.

(22)   What did Caesar drink nectar with?

(23)   *What did Caesar drink nectar and?

   Other examples of what constitutes our linguistic knowledge were mentioned above. The point is that we know more about our language than meets the eye – or ear. How did we come to know so much about the structure and meaning of sentences in our language when what we hear are simply sequences of sounds? This problem of explaining the ease, rapidity and uniformity of language development in the face of impoverished data is called **the logical problem of language acquisition**, and was first posed in this form by Noam Chomsky (1955).

The logical problem of language acquisition (LPLA)

The LPLA is one instance of a more general question concerning the acquisition of human knowledge, which the British philosopher Bertrand Russell summed up as follows:

> How comes it that human beings, whose contacts with the world are brief and personal and limited, are nevertheless able to know as much as we do know? Is the belief in our knowledge partly illusory? And if not, what must we know otherwise than through the senses?

What Russell alludes to in the last line, is the possibility that much of our knowledge is not learned, does not come to us through our experiences with the world, our senses, but is innate, biologically determined.

In the course of acquiring a language, children are exposed to only a finite set of utterances. Yet they come to use and understand an **infinite** set of sentences, as discussed above. This has been referred to as the **creative** aspect of language use. This 'creativity' does not refer to the ability to write poetry or novels but rather the ability to produce and understand an unlimited set of new sentences never spoken or heard previously. The precise linguistic input children receive differs from child to child; no two children are exposed to exactly the same set of utterances. Yet, they all arrive at pretty much the same grammar. The input that children get is haphazard in the sense that caretakers do not talk to their children to illustrate a particular point of grammar. Yet, all children develop systematic knowledge of a language. Thus, despite the severe limitations and variation in the input children receive, and also in their personal circumstances, they all develop a rich and uniform system of linguistic knowledge. The knowledge attained goes beyond the input in various ways. How do we come to know as much as we do about our language if not from the linguistic environment?

In answer to the question of the logical problem of language acquisition, it has been proposed that much of what we know about our language is not in fact learned from the input, but is rather part of an innate endowment, which we referred to above as Universal Grammar (UG). UG specifies the form and functioning of human language in general, hence principles which hold in all languages. On this view, the child's mind does not approach language as a *tabula rasa* (a blank slate) to be written on by experience alone, or armed only with general problem-solving skills such as imitation, memorization, analogy, or general induction. Rather, children are equipped with a set of specific expectations about linguistic structure and the principles which govern language. UG helps them overcome the limitations of the input and guides their grammatical development in

> **SIDEBAR 1.3**
>
> **The problem of no negative evidence**
>
> Ungrammatical sentences such as (23) (repeated here)
>
> (23)   *What did Caesar drink nectar and?
>
> illustrate a particularly thorny aspect of the acquisition problem – the 'no negative evidence problem'. In (23) the questioned element is part of a coordination, as in *Caesar drank nectar and wine*. It seems that there is a restriction against questioning out of a coordinate structure and this restriction holds not just in English, but in all languages. But how can the child learn such a constraint since the relevant evidence is an ungrammatical sentence which he will never hear. The asterisk is a useful linguistic convention, but it is not part of the input to the child. In general, children do not have access to **negative evidence**, that is direct information that certain sentences are ungrammatical – yet another important respect in which the input to the child is deficient.

particular ways. So children develop language rapidly and efficiently, that is, with relatively few errors, and despite the **poverty of the stimulus** (for example, the lack of negative evidence), because the basic form of language is given to them by human biology. Our commonsense understanding of language acquisition is that children just 'pick up' language. This seems to be close to the truth. In fact it may be more appropriate to think in terms of language 'growth' rather than learning. This is not to say that there is no learning. Children must, of course, learn the lexicon of the particular language they are exposed to and other language-specific properties. As we will see throughout this book, languages such as English, French, Japanese, Swahili and so on share many essential properties – those that derive from UG – but they also differ from each other in various respects, and these differences must be learned by the child on the basis of experience. The best way to think about UG is as a kind of template with gaps that must be filled in through experience with a particular language.

There is an intimate connection between linguistic theory and language acquisition. By analyzing the structure of individual languages, linguists try to determine which aspects of our linguistic knowledge are universal and hence, arguably, available to the child as part of UG, and which aspects are language-particular and hence to be learned on the basis of linguistic input that the child receives. Thus, the study of particular grammars tells us something important about language development. At the same time, an understanding of the development of language in children can offer us further insight into the organization of human language. In chapters 6, 10, and 15 we will look more closely at the developmental stages that children go through on their way to adult linguistic competence.

1.6 How is Knowledge of Language Put to Use? Linguistic Performance

If the grammar is the mental representation of linguistic competence, how does this differ from how we use this knowledge when we speak and when we comprehend what is said to us? The distinction between the representation of what we know in our minds and how we put this knowledge to use is not specific to language. In performing music, for example, we may 'know by heart' how to play the *Moonlight Sonata* and we may play it perfectly from time to time. But we may also at some performances, produce a few clinkers, make a number of mistakes, even forget a specific passage. The fact that the next time we try to play it we may make no such mistakes shows that the knowledge was there but we just couldn't get to it in our performance. That is, we know the sonata but our performance reflects this knowledge in a non-perfect way.

This is also true of language. Although in principle we can understand and produce an infinite set of sentences, obviously in our mortal lives no speaker can actually do so. Although on hearing a sentence like (23) above, we know that there is something wrong with it, that it is ungrammatical, thus reflecting our linguistic competence, we may not be able to state what part of our grammar is violated, showing that this is unconscious knowledge.

Differences between linguistic knowledge (competence) and linguistic performance are revealed, for example, through slips of the tongue. When the Reverend Spooner, whose errors gave birth to the term *spoonerism*, referred to Queen Victoria as *that queer old dean* instead of the intended *that dear old queen* he knew that she wasn't a dean and wasn't strange. Everyone makes errors, and often we catch ourselves doing it and correct the errors, showing that we know what is the correct form of the word, phrase, or sentence which is involved in the error.

Memory lapses sometimes prevent us from remembering the beginning of a sentence, producing errors like (24) in which a singular *he* is mistakenly produced instead of the plural *they* to agree with the plural subject Macbeth and Banquo.

(24)   Macbeth and Banquo, two generals of the King, rode to the castle which he saw in the distance.

The relationship between the grammar and linguistic performance is a complex one and is the major area of psycholinguistic research. This text will not discuss linguistic performance but rather the mental system which is accessed in speech and comprehension. However, given this distinction between competence and performance, how can linguists investigate competence, the nature of the mental grammar?

1.7 'Doing' Linguistics

There are a number of methods linguists use in 'doing' linguistics, in analyzing languages, in constructing grammars, and in developing a theory of grammar.

Linguists use both **naturalistic** data and **experimental** data. Naturalistic data consist of actual speech written down or recorded and are often a useful source of **positive evidence**, that is, evidence that a particular type of sentence is (probably) grammatical.

For example, suppose that in listening to recorded conversations or studying pages and pages of transcribed texts, one finds that the pronouns *him* and *her* always refer to someone other than the subject, in sentences like

(25)   Othello hates him.

(26)   Titania loved her.

while in sentences such as (27) and (28),

(27)   Othello hates himself.

(28)   Titania loved herself.

himself and *herself* always refer to the subject; one can infer which forms of the pronouns are grammatical in the different cases. But note the caveat above concerning the difference between competence and performance (specifically between **sentences**, generated by the mental grammar, and **utterances**, what speakers actually say, which, as shown above, may contain 'slips of the tongue,' false starts, the *ers* and *uhs* and *you know's* of filled pauses, etcetera).

In addition, naturalistic data are not a reliable source of negative evidence, that is, evidence that a particular sentence is ungrammatical. This is because it is impossible to infer that a sentence is ungrammatical just because we never hear it. As noted above, the number of possible grammatical sentences in any natural human language is infinite, but people only utter a finite number of sentences in their lives; therefore, many grammatical sentences have never been uttered (much less observed by a linguist). In order to gather negative evidence that a sentence is ungrammatical, linguists generally rely on experimental data, to determine whether native speakers can understand it and to determine whether they find it intuitively natural or acceptable.

In one kind of experiment, native speakers are presented with sentences and asked whether each 'sounds OK,' or whether they think it is an

acceptable sentence. Similar kinds of experiments are conducted in morphology, with native speakers of the language under investigation asked to accept or reject conjoined morphemes, or in phonology, where speakers are asked whether some string of sounds is a possible word in the language. The study of speech sounds – their production, perception, and their acoustic characteristics – is also carried on in the phonetics laboratory with a wide variety of computerized and other technologies.

It is important to bear in mind that the notions of **acceptability** and **grammaticality** are not exactly the same, since factors not directly related to the grammar itself can also affect acceptability. For example, a native speaker might consider a sentence unacceptable because it contains a word that the speaker finds obscene or offensive, or because the sentence is so long they forget how it began, or because the sentence is ambiguous in a way that might be confusing. In such cases, the speaker's judgments about acceptability may not be a fully reliable indicator of grammaticality. Thus, native speakers' intuitions provide only indirect evidence about whether a sentence is grammatical. Since there is no more direct way of finding out whether a sentence is grammatical, linguists generally rely on speakers' intuitions about acceptability, despite these limitations.

To account for judgments such as those in the examples above, we postulate the existence of some form of syntactic rules in the unconscious mental grammar. The hypotheses about particular rules may turn out to be mistaken. If, on further investigation, this proves to be so, new hypotheses embodying the existence of other syntactic rules will have to be formulated to explain speakers' linguistic behavior in both naturalistic and experimental situations.

There will be further discussion on experimental evidence, and on grammaticality judgments and acceptability, in chapter 3. Throughout the book the methods used by linguists in 'doing' linguistics are discussed and you will be 'doing' your own analyses through the many exercises in the chapters that follow.

1.8   Summary

Linguistics is the scientific study of human language. There are many subfields of linguistics. This textbook will be primarily concerned with contemporary linguistic theory, or theoretical linguistics, which aims at an explanation of the nature and structure of all human languages. The interest in human language goes back as far as recorded history. The publication of Chomsky's *Syntactic Structures* in 1957 ushered in the current period of generative linguistics, the aims of which concern answers to

three key questions: what constitutes knowledge of language (linguistic **competence**), how is the knowledge acquired, and how is this knowledge put to use in linguistic **performance**? Speakers' knowledge of their language is represented in their mind as a mental **grammar** which includes everything they know about the language, including its **lexicon** (the words in their vocabulary), the **morphology** (the structure of these words), the **syntax** (the structure of phrases and sentences), the **semantics** (the meanings of linguistic structures), and the **phonetics** and **phonology** (the sounds and sound patterns of the language). Thus a speaker's grammar is the mental representation of linguistic competence. Grammars of individual languages written by linguists, which mirror the mental grammars, are called **descriptive grammars**, which often are used as the basis for **teaching** or **pedagogic grammars** used by those wishing to learn a second (or third) language. Neither mental grammars nor descriptive grammars aim at prescribing what a speaker should say, but rather have as their goal to mirror what speakers know about their language.

Children acquire the complex grammars of their language quickly and with incredible ease without any instruction. How they do this is referred to as the logical problem of child language acquisition. To understand how they do this is a problem since children receive an impoverished set of finite data with no negative evidence (telling them what strings of sounds or words are incorrect) and yet are able to form the rules to produce and comprehend an infinite set of sentences and to distinguish between well-formed (**grammatical**) and ill-formed (**ungrammatical**) utterances. This has led to the suggestion that the human animal is endowed with an innate capacity for language, a **Universal Grammar – UG**.

In trying to understand the nature of language one must distinguish between linguistic **competence** (the representation of linguistic knowledge) and linguistic **performance** (the use of this knowledge in **speech processing** – speaking and comprehending). In performance, our knowledge of language interacts with non-linguistic knowledge and is influenced by many factors including short-term memory, other psychological processes, and pragmatic knowledge, among other things.

Linguists use both **naturalistic** data – actual speech, conversations, recorded texts – and **experimental** data. The chapters that follow will illustrate the methods used by linguists in developing a theory of language and in describing the languages of the world.

Further Reading

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

——. 1968. *Language and Mind*. New York: Harcourt Brace Jovanovich. Extended edition, 1972.

——. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Fromkin, Victoria A. and Robert Rodman. 1998. *An Introduction to Language*. Fort Worth, TX: Harcourt Brace.

Klima, Edward S. and Ursula Bellugi. 1979. *The Signs of Language*. Cambridge, MA: Harvard University Press.

Newmeyer, Frederick J. 1983. *Grammatical Theory: Its Limits and Its Possibilities*. Chicago and London: University of Chicago Press.

Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Languages*. New York: William Morrow and Co.