Fifty years ago the study of language was largely a descriptive endeavor, grounded in the traditions of 19th century European philology. The object of study, as proposed by de Saussure in a famous course of lectures at the University of Geneva (1906–1911), was langue, language as a system, a cultural institution, rather than parole, language as spoken and heard by individuals. In 1933 historical linguists were describing and comparing the world's languages, tracing their family relations, and reconstructing the protolanguages from which they had sprung (Lehmann, 1973). Structural linguists were developing objective procedures for analyzing the sound patterns and syntax of a language, according to well-defined, systematic principles (e.g., Bloomfield, 1933). Students of dialect were applying such procedures to construct atlases of dialect geography (Kurath, 1939), while anthropological linguists were applying them to American Indian, African, Asian, Polynesian and many other languages (Lehmann, 1973). The work still goes on. From it we are coming to understand the origins of language diversity: not only how languages change over time and space but also how they and their dialects act as forces of social cohesion and differentiation (e.g., Labov, 1972).

However, the unfolding of the descriptive tradition and the development of new methods and theories in the field of sociolinguistics are not my concerns in this chapter. My concern, rather, is with a view of language that has emerged from a more diverse tradition. For like the taxonomic studies of Linnaeus in botany and of his followers in zoology, the great labor of language description and classification has provided the raw material for a broader science, stemming from the work of seventeenth century grammarians and of such nineteenth century figures as the German physicist Hermann von Helmholtz, the French neurologist Paul Broca, and the English phonetician Henry Sweet. The several strands that their works represent have come together over the past 30 to 40 years to form the basis of a new science of language, focusing on the individual, rather than on the social and cultural, linguistic system. Since the new focus is essentially biological, a biological analogy may be helpful. It is as though we shifted from describing and classifying the distinctive flight patterns of the world's eight or nine thousand species of birds to analyzing the basic principles of individual

†Also Queens College and Graduate Center, City University of New York.
Acknowledgment. This work was supported in part by NICHD Grant HD-01994 to Haskins Laboratories. I thank Ignatius Mattingly for comments and advice.

[HASKINS LABORATORIES: Status Report on Speech Research SR-84 (1985)]
flight as they must be instantiated in the anatomy and physiology of every hummingbird and condor. Thus, this new science of language asks: What is language as a category of individual behavior? How does it differ from other systems of animal communication? What do individuals know when they know a language? What cognitive, perceptual and motor capacities must they have, to speak, hear, and understand a language? How do these capacities derive from their biophysical structures, that is, from human anatomy and physiology? What is the course of their ontogenetic development? And so on.

Such questions hardly fall within the province of a single discipline. The new field is markedly interdisciplinary, and addresses questions of practical application as readily as questions of pure theory or knowledge. Linguistics, anthropology, psychology, biology, neuropsychology, neurology, and communications engineering all contribute to the field, and their research has implications for workers in many areas of social import: doctors and therapists treating stroke victims, surgeons operating on the brain, applied engineers working on human-machine communication, teachers of second languages, of reading, and of the deaf and otherwise language-handicapped.

The origins of the new science are an object lesson in the interplay between basic and applied research, and between research and theory. To understand this, we must begin by briefly examining the nature of language and the properties that make it unique as a system of communication.

The Structure of Language

If we compare language with other animal communication systems, we are struck by its breadth of reference. The signals of other animals form a closed set with specific, invariant meanings (Wilson, 1975). The ultrasonic squeaks of a young lemming denote alarm; the swinging steps and lifted tail of the male baboon summon his troop to follow; the "song" of the male white-crowned sparrow informs his fellows of his species, sex, local origin, personal identity and readiness to breed or fight. Even the elaborate "dance" of the honey bee merely conveys information about the direction, distance, and quality of a nectar trove. But language can convey information about many more matters than these. In fact, it is the peculiar property of language to set no limit on the meanings it can carry.

How does language achieve this openness, or productivity? There are several key features to its design (Hockett, 1960). Here we note two. First, language is learned: it develops under the control of an open rather than a closed genetic program (Mayr, 1974). Transmission of the code from one generation to the next is therefore discontinuous: Each individual recreates the system for himself. There is ample room here for creative variation—probably a central factor in the evolution of language and in the constant processes of change that all languages undergo (e.g., Kiparsky, 1968; Locke, 1983; Slobin, 1980). One incidental consequence of this freedom is that the universal properties of language (whatever they may be) are largely masked by the surface variety of the several thousand languages, and their many dialects, now spoken in the world.

Second, and more crucially, language has two hierarchically related levels of structure. One level, that of sound pattern, permits the growth of a large lexicon; the other level, that of syntax, permits the formation of an infinitely large set of utterances. A similar combinatorial principle underlies the structure of both levels.
Consider, first, the fact that a 6-year-old, middle-class American child typically has a recognition vocabulary of some 8,000 root words, some 14,000 words in all (Templin, 1957). Most of these have been learned in the previous four years, at a rate of about five or six roots a day. As an adult, the child may come to have a vocabulary of well over 150,000 words (Seashore & Erickson, 1940). How is it possible to produce and perceive so many distinct signals?

The achievement evidently rests on the evolution in our hominid ancestors of a combinatorial principle by which a small set of meaningless elements (phonemes, or consonants and vowels) is repeatedly sampled, and the samples permuted, to form a very large set of meaningful elements (morphemes, words). Most languages have between 20 and 100 phonemes; English has about 40, depending on dialect. The phonemes themselves are formed from an even smaller set of movements, or gestures, made by jaw, lips, tongue, velum, and larynx. Thus, the combinatorial principle was a biologically unique development that provided "a kind of impedance match between an open-ended set of meaningful symbols and a decidedlly limited set of signaling devices" (Studdert-Kennedy & Lane, 1980; cf. Cooper, 1972; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). We may note, incidentally, that a large lexicon is not peculiar to complex, literate societies: Even so-called primitive human groups may deploy a considerable lexicon. For example, the Hanunoo, a stone-age people of the Philippines, have nearly three thousand words for the flora and fauna of their world (Levi-Strauss, 1966).

Of course, a large lexicon is not a language. Many languages have relatively small lexicons, and in everyday speech we may draw habitually on no more than a few thousand words (Miller, 1951). To put words to linguistic use, we must combine them in particular ways. Every language has a set of rules and devices, its syntax, for grouping words into phrases, clauses, and sentences. Among the various devices that a language may use for predicing properties of objects and events, and for specifying their relations (who does what to whom) are word order, and inflection (case, gender, and number affixes for nouns, pronouns, adjectives; person, tense, mood, and voice affixes for verbs). An important distinction is also made in all languages between open-class words with distinct meanings (nouns, verbs, adjectives, etc.) and closed-class or function words (conjunctions, articles, verbal auxiliaries, enclitics) that have no fixed meaning in themselves, but serve the purely syntactic function of indicating relations between words in a sentence or sequence of sentences. Here again then, a combinatorial principle is invoked: a finite set of rules and devices is repeatedly sampled and applied to produce an infinite set of utterances.

I should note that many of the facts about language summarily described above are already framed from the new viewpoint that has developed in the past 40 years. Let us now turn back the clock and consider the early vicissitudes of three areas of applied research that contributed to this development.

Three Areas of Applied Research in Language

In the burst of technological enthusiasm that followed World War II, federal money flowed into three related areas of language study: automatic machine translation, automatic speech recognition, and automatic reading machines for the blind. A considerable research effort was mounted in all three areas during the late 1940s and early 1950s, but surprisingly little
headway was made. The reason for this, as will become clear below, was that all three enterprises were launched under the shield of a behaviorist theory according to which complex behaviors could be properly described as chained sequences of stimuli and responses.

The initial assumption underlying attempts at machine translation was that this task entailed little more than transposing words (or morphemes) from one language into another, following a simple left-to-right sequence. If this were so, we might store a sizable lexicon of matched Russian, say, and English words in a computer and execute translation by instructing the computer to type out the English counterpart of each Russian word typed in. Unfortunately, both semantic and syntactic stumbling blocks lie in the path. The range of meanings, literal and metaphorical, that one language assigns to a word (say, English high, as in "high mountain," "high pitch," "high hopes," "high horse," "high-stepping," and "high on drugs") may be quite different from the range assigned by another language; and the particular meaning to be assigned will be determined by context, that is, by meanings already assigned to some, in principle, unspecified sequence of preceding words. Moreover, the syntactic devices for grouping words into phrases, phrases into clauses, clauses into sentences may be quite different in different languages. This is strikingly obvious when we compare a heavily inflected language, such as Russian, with a lightly inflected language with a more rigid word order, such as English. Oettinger (1972) amusingly illustrates the general difficulties with two simple sentences, immediately intelligible to an English speaker, but a source of knotty problems in both phrase structure and word meaning to a computer, programmed for left-to-right lexical assignment: Time flies like an arrow, and Fruit flies like a banana. From such observations, it gradually became clear that we would make little progress in machine translation without a deeper understanding of syntax and of its relation to meaning.

The initial assumption underlying attempts at automatic speech recognition was similar to that for machine translation and equally in error (cf. Reddy, 1975). The assumption was that the task entailed little more than specifying the invariant acoustic properties associated with each consonant and vowel, in a simple left-to-right sequence. One would then construct an acoustic filter to pass those properties but no others, and control the appropriate key on a printer by means of the output from each filter. Unfortunately, stumbling blocks lie in this path also. A large body of research has demonstrated that speech is not a simple left-to-right sequence of discrete and invariant alphabetic segments, such as we see on a printed page (e.g., Fant, 1962; Joos, 1948; Liberman et al., 1967). The reason for this, as we shall see shortly, is that we do not speak phoneme by phoneme, or even syllable by syllable. At each instant our articulators are engaged in executing patterns of movement that correspond to several neighboring phonemes, including those in neighboring syllables. The result of this shingled pattern of movement is, of course, a shingled pattern of sound. Even more extreme variation may be found when we examine the acoustic structure of the same syllable spoken with different stress or at different rates or by different speakers. From such observations it gradually became clear that we would make little progress in automatic speech recognition without a deeper understanding of how the acoustic structure of the speech signal specifies the linguistic structure of the message.
Finally, the initial assumption underlying attempts to construct a reading machine for the blind was closely related to that for automatic speech recognition and again in error (Cooper, Gaitenby, & Nye, 1984). A reading machine is a device that scans print and uses its contours to control an acoustic signal. It was supposed that, given an adequate device for optical recognition of letters on a page, one need only assign a distinctive auditory pattern to each letter, to be keyed by the optical reader and recorded on tape or played in real time to a listener—a sort of auditory Braille. Once again there were stumbling blocks, but this time they were perceptual. We normally speak and listen to English at a rate of some 150 words per minute (wpm), that is, roughly 5 to 6 syllables or 10 to 15 phonemes per second. Ten to 15 discrete sounds per second is close to the resolving power of the ear (20 elements per second merge perceptually into a low-pitched buzz). Not surprisingly, despite valiant and ingenious attempts to improve the acoustic array, even the most practiced listeners were not able to follow a substitute code at rates much beyond that of skilled Morse code receivers, namely some 10 to 15 words per minute—a rate intolerably slow for any extended use. From this work, it gradually became clear that the only acceptable output from a reading machine would be speech itself. This conclusion was one of many that spurred development of speech synthesis by artificial talking machines in following years (Cooper & Borst, 1952; Fant, 1973; Flanagan, 1983; Mattingly, 1968, 1974). The conclusion also raised theoretical questions. For example: Why can we successfully transpose speech into a visual alphabet, using another sensory modality, if we cannot successfully transpose it within its "natural" modality of sound? Why is speech so much more effective than other acoustic signals? Is there some peculiar, perhaps biologically ordained, relation between speech and the structure of language? We will return to these questions below.

I have not recounted these three failures of applied research missions to argue that money and effort spent on them were wasted. On the contrary, initial failure spurred researchers to revised efforts, and valuable progress has since been made. Reading machines for the blind, using an artificial speech output, have been developed and are already installed in large libraries (Cooper et al., 1984). There now exist automatic speech recognition devices that recognize vocabularies of roughly a thousand words, spoken in limited contexts by a few different speakers (Levinson & Liberman, 1981). Scientific texts with well-defined vocabularies can now be roughly translated by machine, then rendered into acceptable English by an informed human editor.

These advances have largely come about by virtue of brute computational force and technological ingenuity, rather than through real gains in our understanding of language. This is not because we have made no gains, for as we shall see shortly, we surely have. However, none of the devices that speak, listen, or understand actually speaks, listens, or understands according to known principles of human speech and language. For example, a speech synthesizer is the functional equivalent of a human speaker to the extent that it produces intelligible speech. But it obviously does so by quite different means than those that humans use: none of its inorganic components corresponds to the biophysical structures of larynx, tongue, velum, lips, and jaw. Instead, a synthesizer simulates speech by means of a complex system of tuned electronic circuits, and resembles a speaker somewhat as, say, a crane resembles a human lifting a weight. We are still deeply ignorant of the physiological controls by which a speaker precisely coordinates the actions of larynx, tongue, and lips to produce even a single syllable.
Studdert-Kennedy: Some Developments in Research on Language Behavior

In short, the main scientific value of the early work I have described was to reveal the astonishing complexity of speech and language, and the inadequacy of earlier theories to account for it. One important effect of the initial failures was therefore to prepare the ground for a theoretical revolution in linguistics (and psychology) that began to take hold in the late 1950s.

The Generative Revolution in Linguistics

The publication in 1957 of Noam Chomsky's Syntactic Structures began a revolution in linguistics that has been sustained and developed by many subsequent works (e.g., Chomsky, 1965, 1972, 1975, 1980; Chomsky & Halle, 1968). To describe the course of this revolution is well beyond the scope of this chapter. However, the impact of Chomsky's writings on fields outside linguistics--philosophy, psychology, biology, for example--and their importance for the emerging science of language has been so great that some brief exposition of at least their nontechnical aspects is essential. I should emphasize that Chomsky's work has by no means gone unchallenged (e.g., Givon, 1979; Hockett, 1968; Katz, 1981). My intent in what follows is not to present a brief in its defense, but simply to sketch a bare outline of the most influential body of work in modern linguistics.

The central goal of Chomsky's work has been to formalize, with mathematical rigor and precision, the properties of a successful grammar. He defines a grammar as "a device of some sort for producing the sentences of the language under analysis" (Chomsky, 1957, p. 11). A grammar, in Chomsky's view, is not concerned either with the meaning of a sentence or with the physical structures (sounds, script, manual signs) that convey it. The grammar, or syntax, of a language is a purely formal system for arranging the words (or morphemes) of a sentence into a pattern that a native speaker would judge to be grammatically correct or at least acceptable. In Syntactic Structures, Chomsky compared three types of grammar: finite-state, phrase structure, and transformational grammars.

A finite-state grammar generates sentences in a left-to-right fashion: given the first word, each successive word is a function of the immediately preceding word. (Such a model is, of course, precisely that adopted by B. F. Skinner in his Verbal Behavior (1957), a dernier cri in behaviorism, published in the same year as the premier cri of the new linguistics). Chomsky (1956) proved mathematically, as work on machine translation had suggested empirically, that a simple left-to-right grammar can never suffice as the grammar of a natural language. The reason, stated nontechnically, is that there may exist dependencies between words that are not adjacent, and an indefinite number of phrases containing other nonadjacent dependencies may bracket the original pair. Thus, in the sentence, Anyone who eats the fruit is damned, anyone and is damned are interdependent. We can, in principle, continue to add bracketing interdependencies indefinitely, as in Whoever believes that anyone who eats the fruit is damned is wrong, and whoever denies that whoever believes that anyone who eats the fruit is damned is wrong is right.

In practice, we seldom construct such sentences. However, the recursive principle that they illustrate is crucial to every language. The principle permits us to extend our communicative reach by embedding one sentence within another. For example, even a four-year-old child may combine, We picked an
apple and I want an apple for supper into the utterance, I want the apple we picked for supper. Thus, the child embeds an adjectival phrase, we picked (= that we picked with the relative pronoun deleted), to capture two related sentences in a single utterance (cf. Limber, 1973).

Chomsky goes on to consider how we might formulate an alternative and more powerful grammar, based on the traditional constituent analysis of sentences into "parts of speech." Constituent analysis takes advantage of the fact that the words of any language (or an equivalent set of words and affixes) can be grouped into categories (such as noun, pronoun, verb, adjective, adverb, preposition, conjunction, article) and that only certain sequences of these categories form acceptable phrases, clauses, and sentences. By grouping grammatical categories into permissible sequences, we can arrive at what Chomsky terms a phrase-structure grammar. Such a grammar is "a finite set...of initial strings and a finite set...of 'instruction formulas' of the form X→Y interpreted: 'rewrite X as Y'" (Chomsky, 1957, p. 29). Figure 1 illustrates a standard parsing diagram of the utterance, The woman ate the apple, in a form familiar to us from grammar school (above), and as a set of "rewrite rules" from which the parsing diagram can be generated (below).

![Parsing Diagram](image)

**Rewrite Rules**

1. Sentence $\rightarrow$ Noun Phrase + Verb Phrase
2. Noun Phrase $\rightarrow$ Article + Noun
3. Verb Phrase $\rightarrow$ Verb + Noun Phrase
4. Article $\rightarrow$ the, a
5. Noun $\rightarrow$ woman, apple...
6. Verb $\rightarrow$ ate, seized...

Figure 1. Above, a parsing diagram dividing the sentence The woman ate the apple into its constituents. Below, a set of rewrite rules that will generate any sentence having the constituent structure shown above.
Notice, incidentally, that rewrite rules are indifferent to meaning. They will generate anomalous utterances such as The chocolate loved the clock, no less readily than The woman ate the apple. Moreover, many native speakers would be willing to accept such anomalous utterances as grammatically correct, even though they have no meaning. This hints at the possibility that syntactic capacity might be autonomous, a relatively independent component of the language faculty. This is a matter to which we will return below.

An important point about a set of rewrite rules is that it specifies the grouping of words necessary to correct understanding of a sentence. The sentence Let's have some good bread and wine is ambiguous until we know whether the adjective good modifies only bread or both bread and wine. The distinction may seem trivial. But, in fact, the example shows that we are sensitive (or can be made sensitive) to an ambiguity that could not have arisen from any difference in the words themselves or in their sequence. Rather, the origin of the ambiguity lies in our uncertainty as to how the words should be grouped, that is, as to their phrase structure. A correct (or incorrect) interpretation of their meaning therefore depends on the listener (and a fortiori the speaker) being able to assign an abstract phrase structure to the sequence of words.

Whether a complete grammar of English, or any other natural language, could be written as a set of phrase-structure rules is not clear. In any event, Chomsky argues in Syntactic Structures that such a grammar would be unnecessarily repetitive and complex, since it does not capture a native speaker's intuition that certain classes of sentence are structurally related. For example, the active sentence Eve ate the apple and the passive sentence, The apple was eaten by Eve could both be generated by an appropriate set of phrase-structure rules, but the rules would be different for active sentences than for their passive counterparts. Surely, the argument runs, it would be "simpler" if the grammar somehow acknowledged their structural relation by deriving both sentences from a common underlying "deep structure." The derivation would be accomplished by a series of steps or "transformations" whose functions are to delete, modify, or change the order of the base constituents Eve, ate, apple.

An important aspect of transformations is that they are structure dependent, that is, they depend on the analysis of a sentence into its structural components, or constituents. For example, to transform such a declarative sentence as The man is in the garden into its associated interrogative Is the man in the garden?, a simple left-to-right rule would be: "Move the first occurrence of is to the front." However, the rule would not then serve for such a sentence as The man who is tall is in the garden, since it would yield Is the man who tall is in the garden? The rule must therefore be something like: "Find the first occurrence of is following the first noun phrase, and move it to the front" (Chomsky, 1975, pp. 30-31). Thus, a transformational grammar, no less than a phrase-structure grammar, presupposes analysis of an utterance into its grammatical (or phrasal) constituents. We may note, in passing, that children learning a language never produce sentences such as Is the man who tall is in the garden? Rather, their errors suggest that, even in their earliest attempts to frame a complex sentence, they draw on a capacity to recognize the structural components of an utterance.
However, here we should be cautious. Chomsky has repeatedly emphasized that "... a generative grammar is not a model for a speaker or hearer" (1965, p. 9), not a model of psychological processes presumed to be going on as we speak and listen. The word "generative" is perhaps misleading in this regard. Certainly, experimental psychologists during the 1960s devoted much ingenuity and effort to testing the psychological reality of transformations (for reviews, see Cairns & Cairns, 1976; Fodor, Bever, & Garrett, 1974; Foss & Hakes, 1978). But the net outcome of this work was to demonstrate the force of Chomsky's distinction between formal descriptions of a language and the strategies that speakers and listeners deploy in communicating with each other (cf. Bever, 1970).

At first glance, the distinction might seem to be precisely that between language and parole, drawn by de Saussure. However, for de Saussure, language, the system of language, "exists only by virtue of a sort of contract signed by the members of a community" (de Saussure, 1966, p. 14): it is a kind of formal artifice or convention, maintained by social processes of which individuals may be quite unaware. By contrast, for Chomsky the "generative grammar [of a language] attempts to specify what the speaker actually knows" (1965, p. 8). What a speaker knows, competence in Chomsky's terminology, is attested to by "intuitive" judgments of grammaticality. What a speaker does, performance (parole), is linguistic competence filtered through the indecisions, memory lapses, false starts, stammerings, and the "thousand natural [nonlinguistic] shocks that flesh is heir to." Thus, even though a theory of grammar is not a theory of psychological process, it is a theory of individual linguistic capacity.

In Chomsky's view, the task of linguistics is to describe the structure of language much as an anatomist might describe the structure of the human hand. The complementary role of psychology in language research is to describe language function and its course of behavioral development in the individual, while physiology, neurology, and psychoneurology chart its underlying structures and mechanisms.

Whether this sharp distinction between language as a formal object and language as a mode of biological function can, or should, be maintained is an open question. What is clear, however, is that it was from a rigorous analysis of the formal properties of syntax (and, later, of phonology: see Chomsky & Halle, 1968) that Chomsky was led to view language as an autonomous system, distinct from other cognitive systems of the human mind (cf. Fodor, 1982; Pylyshyn, 1980). His writings during the late 1950s and 1960s brought an exhilarating breath of fresh air to psychologists interested in language, because they offered an escape from the stifling behavioristic impasse, already noted by Lashley (1951) and others (e.g., Miller, Galanter, & Pribram, 1960).

The result was an explosion of research in the psychology of language, with a strong emphasis on its biological underpinnings. Whatever one's view of generative grammar, it is fair to say that almost every area of language study over the past 25 years has been touched, directly or indirectly, whether into action or into reaction, by Chomsky's work. This will be obvious from the following selective review of research in four major areas: acoustic phonetics, American Sign Language (ASL), brain specialization for language, and language development in children.
Acoustic Phonetics

We begin with audible speech, partly because we are then following the course of development, both in the species and the individual, from the bottom up; partly because it is in this area, where we are dealing with observable, physical processes, that the most dramatic progress has been made; and partly because we have come to realize in recent years that the physical medium of language places fundamental constraints on its surface structure. To understand this we must know something of the way speech is produced.

The source-filter theory of speech production. The source-filter theory, first proposed by Johannes Müller in 1848, has been elaborated in the past 50 years, notably at the University of Tokyo (Chiba & Kajiyama, 1941), the Royal Institute of Technology in Stockholm (Fant, 1960, 1973) and, in this country, the Massachusetts Institute of Technology (Stevens & House, 1955, 1961) and Bell Telephone Laboratories (Flanagan, 1983). As a result of this work, we are now able to specify accurately the possible acoustic outputs of any vocal tract, animal or human.

When we speak, we drive air from our lungs through the pharynx, mouth, teeth, lips and, sometimes, nose. The sound source is usually either the "voice" produced by rapid pulsing of the vocal cords (as in the final sounds of be and do), the hiss of air blown through a narrow constriction (as in the initial and final sounds of safe and thrush) or both (as in the final sounds of leave and bees). The resonant filter is the vocal tract, its air set into vibration by the flow of air from the lungs, much as we produce sound from a bottle or a wind instrument by blowing air across its top.

To some large degree linguistic information (that is, consonants and vowels) is conveyed by systematic variations in the configuration of the vocal tract. For example, if we lower the tongue and move it back toward the pharynx, we set up a pattern of resonances (known as formants) corresponding to the vowel [a]. If we raise the tongue forward toward the gums, we set up resonances for the vowel [i]. Finally, if we raise the tongue backward toward the soft palate, we set up resonances for the vowel [u]. These three sounds are the most distinct vowels, both articulatorily and acoustically, that the human vocal tract can produce, and all known languages use at least two of them.

(We may note, in passing, that Lieberman and his colleagues [Lieberman & Crelin, 1971; Lieberman et al., 1972]) have used the source-filter theory of speech production to demonstrate that these vowels lie outside the range of sounds that could be produced either by an adult chimpanzee or by a newborn human infant. The reason for this is that the larynx in both chimpanzee and infant is high in the throat, restricting the range of possible tongue movements. An advantage of the high larynx for the infant is that it provides an arrangement of the oral tract such that, like other mammals, the infant can suck through its mouth and breathe through its nose at the same time. Over the first six months of life, the infant's larynx lowers, a special swallowing reflex develops to prevent food entering the lungs, and the infant becomes capable of producing the vowels of the language spoken around it. The lowered larynx seems to be one of several adaptations of the vocal apparatus that have suited it for speaking as well as for eating and breathing.)
Of course, we do not speak only in vowels. Rather, we speak in runs of syllables, alternately constricting the vocal tract to form consonants, opening it to form vowels. (This repeated opening and closing of the tract produces the rises and falls of amplitude that are the basis of speech rhythm and poetic meter.) What is of interest, as we have already remarked, is that the tract configurations appropriate to particular consonants and vowels do not follow each other in linear sequence. At any instant, each articulator is executing a complex pattern of movement, of which the spatiotemporal coordinates reflect the influence of several neighboring segments. Readers may test this by slowly uttering, for example, the words cool and keel. They will find that the position of the tongue on the palate during closure for the initial consonant, [k], is slightly further back for the first word than for the second. The result of this interleaving is that, at any instant, the sound is conveying information about more than one phonetic segment, and that each phonetic segment draws information from more than one piece of sound—an obvious problem for automated speech recognition. Unfortunately, we cannot, as was at one time hoped, escape from this predicament by building a machine to recognize syllables, because similar interactions between phonetic segments occur across syllable boundaries. We see all this quite clearly if we examine a sound spectrogram.

The sound spectrograph. The sound spectrograph was developed at Bell Telephone Laboratories during World War II, to provide a visible display of the acoustic spectrum of speech as it changes over time. Originally, it was hoped that the device would enable deaf persons to use the telephone (Potter, Kopp, & Green, 1947), but this proved impracticable because spectrograms are formidable difficulty to read (though see Cole et al., 1980).

Figure 2 is a spectrogram of the utterance She began to read her book. Frequency on the ordinate is plotted against time on the abscissa. Variations in relative amplitude appear as variations in the darkness of the pattern. The dark bars correspond to formants, that is, to resonant peaks in the vocal tract resonance function. Scattered patches, as at the beginning, correspond to the noise of fricatives, e.g., [f], [s], and stop consonants, e.g., [p], [b]. A series of vertical lines has been drawn, dividing the spectrogram into discrete, acoustic segments. There are 25 of these segments, even though the utterance consists of only 17 phonetic segments and 7 syllables. Some of these acoustic segments correspond more or less directly to phonetic segments: thus, segments 1 and 2 correspond to the two sounds of she. Segment 3, on the other hand, corresponds to the first three sounds of began, segments 11 and 12 to the first sound of to, segment 23 to the first two sounds of book.

The sound spectrograph revealed, for the first time, the astonishing variability of the speech signal both within and across speakers. It was also the basis for the first systematic studies of speech perception, from which we have learned which aspects of the signal carry crucial phonetic information. These studies, in turn, provided the basis for the development of speech synthesis. Thus, artificial talking machines, now being used in reading machines for the blind and in a variety of human-machine communication systems, rest squarely on the shoulders of the spectrograph.

Speech perception. Early work in speech perception was largely guided by the demands of telephonic communication. Its aim was to estimate how much distortion (by filtering, noise, peak-clipping, and so on) could be imposed on the signal without seriously reducing its intelligibility (Licklider & Miller,
Figure 2. A spectrogram of the utterance She began to read her book. Frequency is plotted on the ordinate, time on the abscissa; relative amplitude is represented by varying degrees of darkness in the display. The dark horizontal bands reflect resonant peaks in the vocal tract transfer function (formants, conventionally numbered from the bottom up: first formant, second formant, etc.); the vertical striations reflect repeated opening and closing of the glottis (voice). Heavy vertical lines have been drawn dividing the pattern into 25 discrete acoustic segments (see text).

1951; Miller, 1951). Two general conclusions from this work were surprising and important. First, speech is so resistant to distortion that we can throw away large parts of the signal without reducing its intelligibility. Second, intelligibility does not depend on naturalness. These two facts made it possible to learn a great deal about the important information-bearing elements in speech by stripping it down to its minimal cues.

Work of this kind was first undertaken at Haskins Laboratories in New York during the 1950s, as part of a program to develop a suitable output for a reading machine. The key research tool was the Pattern Playback, developed by F. S. Cooper (Cooper, 1950; Cooper & Borst, 1952) to reconvert the visual pattern of a spectrogram into sound. The pattern, painted on a moving acetate belt, reflects frequency-modulated light to a photocell that drives a speaker. Figure 3 illustrates an early spectrogram and its stylized copy. If the copy is passed through the playback, it produces an intelligible version of the utterance To catch pink salmon. The utterance sounds unnatural, partly because the formant bandwidths have been sharply reduced, partly because it is spoken in a monotone.

The playback made it possible for experimenters to manipulate the speech signal systematically, by pruning, deleting, or exaggerating portions of the spectrographic pattern until they had determined the minimal cues for any
Figure 3. Above, a spectrogram of the utterance To catch pink salmon. Below, a stylized copy of the spectrogram, sufficient to regenerate the utterance if played on the Pattern Playback.
particular utterance (Liberman, 1957; Liberman et al., 1959). With this device, and with its successors at Haskins and elsewhere, a body of knowledge was built up, sufficient for synthesis by rule of relatively high-quality speech (Fant, 1960, 1968; Flanagan, 1983; Mattingly, 1974).

Several reviews of the perceptual implications of this work have been published (Darwin, 1976; Liberman et al., 1967; Liberman & Studdert-Kennedy, 1978; Studdert-Kennedy, 1974, 1976), and I will not review them here. However, two facts deserve note. First, the cues for a given phonetic segment (that is, for a particular consonant or vowel) vary markedly as a function of context. Figure 4 displays spectrograms of the naturally spoken syllables [did] and [dud]. We know from synthetic speech that a main cue to the initial [d] lies in changes in the second formant after onset. Notice that the second formant rises before [i], falls before [u], and that the rising and falling patterns are precisely reversed for the final [d]. Yet all are heard as [d]. Moreover, if these patterns or their synthetic versions are removed from context and presented to listeners for judgments, they are no longer heard as [d], nor are they heard as invariant. Rather they are heard as rising and falling tones (Liberman et al., 1967). In other words, different acoustic patterns are heard as different in a non-speech context but as the same in a speech context. This is merely one of dozens of such examples.

![Spectrograms of naturally spoken syllables](image)

**Figure 4.** Spectrograms of naturally spoken [did] (deed) and [dud] (dood). The acoustic information specifying the alveolar place of articulation of the initial and final consonants is primarily carried by the second formant, centered around 2kHz for [did] and slightly below 1 kHz for [dud]. Note that this formant forms a parabola, concave downwards in [did], concave upwards in [dud]. Despite this difference, both patterns are heard as beginning and ending with [d].
Studdert-Kennedy: Some Developments in Research on Language Behavior

The second fact of note is that despite the apparent lack of discrete phonetic segments in the signal, listeners have little difficulty in learning to find segments—so little, in fact, that a segmental representation of speech is the basis of the alphabet.

The interpretation of these facts is still a matter of controversy (e.g., Cole & Scott, 1974; Ladefoged, 1980; Stevens, 1975), and I will not pursue the matter here. However, it is worth noting that such findings gave rise to the hypothesis that humans have evolved a specialized perceptual mechanism for speech, distinct from, though dependent on, their general auditory system (Liberman, 1970, 1982; Liberman et al., 1967; Liberman & Studdert-Kennedy, 1978). The hypothesis has received substantial support from many dozens of studies of dichotic listening over the past 20 years (e.g., Kimura, 1961, 1967; Shankweiler & Studdert-Kennedy, 1967; Studdert-Kennedy & Shankweiler, 1970; for a review, see Porter & Hughes, 1983). The conclusion from this work, and from studies of patients with separated cerebral hemispheres (see section below on brain specialization for language), is that the left hemisphere of most normal right-handed individuals is specialized not only for speaking (as has been known for many years from studies of brain-damaged patients), but also for perceiving speech. Specifically, there is now good reason to believe that "while the general auditory system common to both hemispheres is equipped to extract the auditory parameters of a speech signal, the dominant [i.e., left] hemisphere may be specialized for the extraction of linguistic features from these parameters" (Studdert-Kennedy & Shankweiler, 1970, p. 579).

An important implication of this conclusion is that speech forms an integral part of the left-hemisphere language system discussed below. With this in mind let us turn to recent work on American Sign Language, which draws on a different perceptucomotor system than spoken language.

American Sign Language

Speech is the natural medium of language. Specialized structures and functions have evolved for spoken communication: vocal tract morphology, lip, jaw, and tongue innervation, mechanisms of breath control (Lenneberg, 1967), and perhaps even (as I have just suggested) matching perceptual mechanisms. But is there any further specialization for language? Is language an autonomous system, distinct from other cognitive systems, as Chomsky has argued?

An opportunity to address this question has arisen in recent years from an unexpected quarter: sign languages of the deaf. Until some 20 years ago, it was commonly believed that sign languages of the deaf—and of other social groups, such as American Plains Indians and Australian aborigines—were either more or less impoverished hybrids of conventional iconic gesture and impromptu pantomime, or artificial systems based, like reading and writing, on a specific spoken language. Artificial systems, such as Signed English and Paget-Gorman, are indeed used in many schools of the deaf: their signs refer to letters (finger-spelling) or higher-order linguistic units (words, morphemes), and their syntax follows that of the base language. However, there are other signed languages, not based on any spoken language, with their own independent lexicons and syntactic systems. The most extensively studied of these is American Sign Language (ASL), the first language of over 100,000 deaf individuals and, according to Mayberry (1978), the fourth most common language (after English, Spanish, and Italian) in the United States.
Modern ASL stems from a French-based sign language introduced into the United States by Thomas Gallaudet in 1817. (According to Stokoe [1974] ASL signers today find French SL more intelligible than British SL, a nice demonstration that ASL is independent of English.) Thus, the original language was in fact based on a spoken language. However, over the past 165 years it has developed among the deaf into an independent sign language.

Structural analysis of ASL was first undertaken by Stokoe (1960), and in 1965 he and his colleagues (Stokoe, Casterline, & Croneberg, 1965) published A Dictionary of American Sign Language on Linguistic Principles, containing a description and English gloss of nearly 2500 signs. The dictionary used minimal pair analysis to show that signs contrasted along three independent dimensions: hand configuration, place of articulation, and movement. For example, signs for APPLE and JEALOUS contrast in hand configuration; signs for SUMMER and UGLY contrast in place of articulation; signs for CHAIR and TRAIN contrast in movement (Klima & Bellugi, 1979, p. 42). Stokoe et al. isolated 55 "chereomes" or primes, analogous to the phonemes of a spoken language: 19 for hand configuration, 12 for place of articulation, and 24 for movement. Thus, they demonstrated that ASL has a sublexical structure, analogous to the phonological structure of a spoken language.

ASL also has a second level of structure, a grammar or syntax. This has been demonstrated in an extensive program of research at the Salk Institute for Biological Studies in La Jolla, over the past 10 years (Klima & Bellugi, 1979). I will not attempt to review this work in any detail, but several points deserve note. First, ASL has a rule-governed system of compounding, by which signs may be combined to form a new sign different in meaning from its components. The process is analogous to that by which, in English, hard and hat, say, are combined to form hardhat, meaning a construction worker. Thus, the lexicon of ASL can be expanded by rule, not simply by iconic invention.

Second, ASL has an elaborate system of inflections by which it modulates the meaning of a word. For example, in English, changes in aspeclual meaning (that is, distinctions in the onset, duration, frequency, recurrence, permanence, or intensity of an event) are indicated by concatenating morphemes. We may say, he is quiet, he became quiet, he used to be quiet, he tends to be quiet, and so on. All these meanings are conveyed in ASL by distinct modulations of the root sign's movement. In the root sign for QUIET the hands move straight down from the mouth, while for TENDS TO BE QUIET they move down forming a circle. Similarly, related nouns and verbs are also distinguished by movements, while verbs are inflected by movement modulation for person, number, reciprocal action, and aspect.

Third, ASL has a spatial (rather than a temporal) syntax. Nouns introduced into a discourse are assigned arbitrary reference points in a horizontal plane in front of the signer. These points then serve to index grammatical relations among referents: verb signs are executed with a movement between two points, or across several points, to indicate subject and object. Thus, a grammatical function variously served in spoken language by word order, case markers, verb inflections, and pronouns is fulfilled in ASL by a spatial device.

Finally, ASL has a variety of syntactic devices that make use of the face. Liddell (1978) has shown that a relative clause ("The apple that Eve offered tempted him") may be marked by tilting back the head, raising the
eyebrows, and tensing the upper lip for the duration of the clause. Baker and Padden (1978) describe gestures of the face and head that mark the juncture of conditional clauses ("If you eat the fruit, you will be punished").

In short, though structural analysis of ASL is far from complete, it is evident that the language has a dual pattern of form and syntax, fully analogous to that of a spoken language. Nonetheless, there are differences. The main structural difference between ASL and English was illustrated by Klima and Bellugi (1979) in a comparison of their rates of communication. The times taken to tell a story in the two languages were almost exactly equal. Yet the speaker used two to three times as many words as the signer used signs. The reason for the discrepancy, already hinted at, lies in the temporal distribution of information. Speech, for the most part, develops its patterns in time, sequentially, while ASL develops its patterns both simultaneously, in space, and sequentially. The difference is evidently due to the difference in the perceptual modalities addressed. Sign, addressed to the eye, is free to package information in parallel; speech, addressed to the ear, is forced into a serial mode. What is interesting, of course, is that despite constraints of modality, the two languages convey information at roughly the same rate. This suggests that they may be operating under the same temporal constraints of cognition.

What, finally, are the implications of this work for the study of speech and language? Evidently, the dual structure of language is not a mere consequence of perceptuomotor modality, but a reflection of cognitive requirements. Whether these cognitive requirements are linguistic rather than general is still not clear. Differently put, we still do not know whether the relation between signed and spoken language is one of analogy or homology. If the two systems prove to be homologous, that is, if they prove to draw on the same neural structures and organization, we will have strong evidence that language is a distinct cognitive faculty. However, if they do not draw on the same underlying neural organization, we might suppose that linguistic structure is purely functional, the adventitious consequence of a cognitively complex animal's attempt to communicate its thought. Studies of sign-language breakdown due to brain injury, discussed below, are therefore of unusual interest and importance.

Brain Specialization for Language

Most of our knowledge of brain specialization for language comes from those "experiments of nature" in which some more or less circumscribed lesion (due to stroke, epilepsy, congenital malformation, gunshot wounds, and so on) proves to be correlated with some more or less circumscribed cognitive or linguistic deficit (for a brief account of modern brain-scanning techniques, see Benson, 1983, and references therein). Recently, our sources of knowledge have been expanded by use of brain stimulation, preparatory to surgery under local anesthesia (Ojemann, 1983, and references therein), and by studies of so-called "split-brain" patients whose cerebral hemispheres have been separated surgically for relief of epilepsy (see below). Some degree of concordance between patterns of brain localization in normal and abnormal individuals has been established by experiments on normals in which visual or auditory input is confined, or more clearly delivered, to one hemisphere rather than the other (Moscovitch, 1983).
Evidence from studies of aphasia. The term aphasia refers to some impairment in language function, whether of comprehension, production, or both, due to some more or less well-localized damage to the brain. Systematic study of aphasia goes back well over a hundred years, and the literature of the subject is vast (for reviews, see, for example, Goodglass & Geschwind, 1976; Heiden & Albert, 1978; Lesser, 1978; Luria, 1966, 1970). The most that can be done here is to hint at one area in which linguistics (that is, formal language description) has begun to affect aphasia studies.

Until recently, the standard framework for describing aphasic symptoms was that of the language modalities: speaking, listening, reading, and writing, or, more generally, the dimensions of expression and reception. These are still the dimensions of the major test batteries used to diagnose aphasia, such as the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1972). An important assumption, underlying any attempt at diagnosis, is that damage to a particular region of the brain has particular, not general, effects on language function. The assumption has strong empirical support and has led to the isolation of two (among several other) broad types of aphasia, nonfluent and fluent, respectively associated with damage to the left cerebral hemisphere in an anterior region around the third frontal convolution (Broca's area) and a posterior region around the superior temporal convolution (Wernicke's area).

Broca's area lies close to the motor strip of the cortex (in fact, close to that portion of the strip associated with motor control of the jaw, lips, and tongue), while Wernicke's area surrounds the primary auditory region. In accord with this anatomical dissociation, a Broca's aphasic (that is, an individual with damage to Broca's area) has been classically found to be nonfluent: having good comprehension, but awkward speech, characterized by pauses, difficulties in word-finding, and distorted articulation; utterances are described as "telegraphic," consisting of simple, declarative sentences, relying on nouns and uninflected verbs, omitting grammatical morphemes or function words. By contrast, a Wernicke's aphasic has been found to have poor comprehension, even of single words, but fluent speech, composed of inappropriate or nonexistent (though phonologically correct) words, often inappropriately inflected and/or out of order.

Notice that these descriptions are still couched in terms of input and output—that is, modalities of behavior—rather than in linguistic terms. The idea that linguistic theory should be brought to bear on aphasia, and attempts made to characterize deficits in terms of overarching linguistic function, has been proposed a number of times in the past (e.g., Jakobson, 1941; Pick, 1913). But only recently (again, partly under the influence of Chomsky's view of language as an autonomous system, composed of autonomous syntactic and phonological subsystems) has the idea begun to receive widespread attention. The general hypothesis of the studies described below is that language breaks down along linguistic rather than modal lines of demarcation.

We will focus mainly on the hypothesis that syntactic competence is discretely and coherently represented in Broca's area of the left frontal lobe. If this is so, the clinical impression that Broca's aphasics have good comprehension, despite their agrammatic speech (and, incidentally, writing), must be in error. More careful testing should reveal deficits in their comprehension, also.
Caramazza and Zurif (1976) tested this hypothesis with three types of sentence: (1) simple declarative sentences in which semantic constraints might permit decoding without appeal to syntax (The apple that the boy is eating is red); (2) so-called reversible sentences that require knowledge of syntactic relations for decoding (The boy that the girl is chasing is tall); and (3) implausible, though grammatically correct, sentences (The boy that the dog is patting is fat). The sentences were presented orally and patients were asked to choose which of two pictures represented the meaning of the sentence. The incorrect alternative showed either a subject-object reversal or an action different from that specified by the verb.

Broca's aphasics performed very well on simple declarative sentences and on sentences with strong semantic constraints (as when the incorrect alternative depicted the wrong action). On reversible plausible and implausible sentences (when the incorrect alternative depicted a subject-object reversal) the patients' performance was at chance. Caramazza and Zurif (1976) concluded that the clinical impression of good comprehension in Broca's aphasics was due to their ability to draw on semantic and pragmatic constraints to understand sentences despite their inability to process syntax.

Other studies have shown that Broca's aphasics a) have difficulty in parsing a sentence into its grammatical constituents (von Stockert, 1972); b) cannot use articles to assign appropriate reference in understanding a sentence (Goodenough, Zurif & Weintraub, 1977), and c) cannot, in general, access closed-class grammatical morphemes (Zurif & Blumstein, 1978). These studies are not without their critics (e.g., Linebarger, Schwartz, & Saffran, 1983), nor is the general claim that aphasic breakdown is typically (or, indeed, ever) along purely linguistic lines (Studdert-Kennedy, 1983, pp. 193-194): the locus and extent of brain damage in aphasia is largely a matter of chance, and it is rare that language alone is affected. However, we have other sources of evidence to test the hypothesis that syntax is represented in the brain as a functionally discrete subsystem.

Evidence from split-brain studies. One source of evidence is the split-brain patient whose cerebral hemispheres have been separated surgically for relief of epilepsy. The condition permits an investigator to assess the cognitive and linguistic capacities of each hemisphere separately. Zaidel (1978) has devised a contact lens, opaque on either the nasal or temporal side, that can be used (profiting from decussation of the optic pathways) to ensure that visual information is freely scanned by a single hemisphere. A variety of written verbal materials--nonsense syllables, words, sentences of varying length and complexity--and pictures can then be used to test the capacities of the isolated hemispheres. For example, the sentences, The fish is eating or The fish are eating, can be presented to a single hemisphere, together with appropriate alternative pictures, to test the hemisphere's capacity to understand written verbal auxiliaries (is, are) (Zaidel, 1983). Similarly, pictures of various objects belonging to different classes (fruit, furniture, vehicles, etc.) might be presented to a single hemisphere to test the hemisphere's capacity to categorize.

The number of available subjects is, of course, limited. But the conclusions from studies of four split-brain patients are remarkably consistent (Zaidel, 1978, 1980, 1983). In general, each hemisphere seems to have "a complete cognitive system with its own perception, memory, language, and cognitive abilities, but with a unique profile of competencies: good on some abilities, poor on others" (Zaidel, 1980, p. 318). Of particular
interest in the present context is the finding that, although the right hemisphere cannot speak, it has a sizable auditory and reading lexicon. However, unlike the left hemisphere, the right cannot read new (nonsense or unknown) words or recognize words for which it has no semantic interpretation. Similarly, the right hemisphere cannot group pictures of objects on the basis of rhyme (e.g., nail, male). Evidently, phonological analysis is the prerogative of the left hemisphere.

The syntactic capacity of the right hemisphere is also limited. The hemisphere can recognize verbal auxiliaries (see above), but has difficulty in discriminating inflections (The fish eat vs. The fish eats). Similarly, the right hemisphere can recognize and interpret nouns, adjectives, and certain prepositions, but has difficulty with the English infinitive marker to. These findings on closed-class morphemes mesh to a degree with the deficits of Broca's aphasics, described above. Not surprisingly, the right hemisphere's capacity to understand sentences is sharply reduced: it cannot deal with sentences longer than about three words.

On the evidence of these studies, then, the right hemisphere has essentially no phonological capacity and only a limited syntactic capacity. Unfortunately, the limited syntactic capacity is equivocal because all these split-brain patients have had epilepsy since early childhood. Brain disorders are known to lead to reorganization and redistribution of function, particularly in childhood (Lenneberg, 1967; Dennis, 1983). We cannot therefore be sure that such syntactic capacity as the right hemisphere displays does not reflect compensation for left hemisphere deficiencies, induced by epilepsy.

Evidence from studies of ASL "aphasia." Studies of normally hearing, brain-damaged patients have established a double dissociation of brain locus and function in right-handed individuals: the left cerebral hemisphere is specialized for language, the right hemisphere for visual-spatial functions (as revealed, for example, by tests requiring a subject to copy a drawing, assemble wooden blocks into a pattern, or discriminate between photographs of unfamiliar faces). As we have seen, ASL is an autonomous linguistic system with a dual structure analogous to that of spoken language, on the one hand, yet, on the other, it encodes its meanings in visual-spatial rather than auditory-temporal patterns. How then should we expect brain damage to affect the language of a native ASL signer?

The answer bears directly on our understanding of the basis of brain specialization for language. For if language loss in ASL aphasia follows damage to the right hemisphere, we may infer that language is drawn to the hemisphere controlling its perceptuomotor channel of communication. But if language loss follows damage to the left hemisphere, we may infer that the neural structure of that hemisphere is, in some sense, matched to the structure of language, whatever its modality. Language might then be seen as a distinct cognitive faculty, sufficiently abstract in its descriptive predicates to encompass both speaking and signing.

Recent studies at the Salk Institute, the first systematic and linguistically motivated studies of ASL aphasia on record, support the second hypothesis. Moreover, the forms of ASL breakdown vary with locus of lesion in a fashion strikingly similar to certain forms of spoken-language breakdown. Bellugi, Poizner, and Klima (1983) describe three patients, all of whom are native ASL signers and display normal visual-spatial capacity for nonlanguage
functions. Their symptoms, resulting from strokes, divide readily into the
two broad classes noted above for spoken language: two patients are fluent,
one is nonfluent.

The two fluent patients display quite different symptoms, coordinated
with different areas of damage to the left hemisphere. The deficits of one
patient (PD) are primarily grammatical; the deficits of the other (KL) are
primarily lexical. PD has extensive subcortical damage from below Broca's
area in the frontal lobe through the parietal to the temporal lobe, abutting
Wernicke's area. PD produces basically normal root signs, but displays an
abundance of semantic and grammatical paraphasias. He produces many
semantically displaced signs (e.g., EARTH for ROOM, BED for CHAIR, DAUGHTER
for WIFE). More strikingly, he often modulates an appropriate root form with
an inappropriate or nonsensical inflection. Finally (despite his normal,
nonlanguage visual-spatial capacity), his spatial syntax is severely
disordered: he misuses or avoids spatial indexing (the equivalent of
pronominal function, as noted above), and overuses nouns.

The second fluent patient, KL, has more limited damage, extending in a
strip across the left parietal lobe. Her deficits, though relatively mild,
are almost the reverse of PD's. First, she avoids nouns and overuses pronouns
(spatial indexing). Second, she tends to make formational errors in root
signs, producing nonsense items by substituting incorrect hand configurations,
places of articulation, or movements. Thus, these two fluent patients display
almost complementary deficits, breaking along linguistic fault lines, as it
were, between lexicon and grammar.

The third patient (CD) is nonfluent. She has massive damage over most of
the left frontal lobe, including Broca's area. She produces individual signs
correctly (with her nondominant hand, due to paralysis of the right side of
her body), and can repeat a test series of signs rapidly and accurately, so
that her deficits are not simply motoric. Yet her spontaneous signing invites
description by just those epithets that characterize a Broca's aphasic. Her
utterances are slow, effortful, short, and agrammatic, largely made up of
open-class items. She omits all grammatical formatives, including
inflections, morphological modulations, and most spatial indices. In short,
this patient, too, displays a peculiarly linguistic rather than a general
cognitive pattern of breakdown.

From this brief review of brain specialization for language we may draw
several conclusions. First, language breakdown seems to follow rough
linguistic lines of demarcation, indicating that phonology (or patterns of
sign formation) and syntax may be supported by separable neural subsystems
within the left hemisphere. Second, left hemisphere specialization does not
rest on a particular sensorimotor channel. Rather, the hemisphere supports
general linguistic functions, common to both spoken and signed language.
Thus, despite the left hemisphere's innate predisposition for speech (see
below on language acquisition), its initial neural organization is
sufficiently plastic to admit quite different language forms (cf. Neville,
1980; Neville, Kutas, & Schmidt, 1982). At the same time, we still do not
know enough about the anatomy and physiology of the brain to be sure that
areas important for particular functions in spoken language precisely
correspond to areas important for analogous functions in signed language: the
issue of analogy vs. homology is not yet closed.
Several further cautions should be noted. It is not yet clear (either from linguistic theory or from behavioral evidence) that syntax and phonology constitute homogeneous functions: some aspects of syntax and phonology may be separable from some aspects of language, others may not (Dennis, 1983). Second, it is even less clear that we should expect a coherent function, once specified, to be discretely and coherently localized in the brain. In looking for correspondences between one level of description (linguistic) and another level (neurological), we may be guilty of the "first-order isomorphism fallacy" that caused the downfall of phrenology and faculty psychology. The error would be analogous to that of someone who expected a single function of an automobile—say, acceleration—to be discretely and coherently localized in the engine. In fact, of course, the mechanism underlying acceleration is distributed over gears, fuel pump, carburetor, pistons, and so on. Perhaps syntactic and phonological functions emerge, like acceleration, from the coordinated actions of disparate parts.

Language Acquisition

As many as 5 percent of American children suffer from some form of delayed or disordered language development, and many more join the ranks of the illiterate. Moreover, there is growing evidence that the capacity to read depends in large part on normal development of the primary language processes of speaking and listening (Crain & Shankweiler, in press). Scientific understanding of development is therefore of broad pediatric and educational interest. In the first instance, the work may simply permit us to establish reliable norms, based on a sound understanding of what language acquisition entails. Later, we may hope, the work should lead to more effective therapeutic intervention than is now available.

No area of language study has been more strongly affected by Chomsky's work than language acquisition. Indeed, it is fair to say that until Chomsky's writings began to be widely disseminated among psychologists, in the early 1960s, the field did not exist. The few psychologists who considered the matter at all (e.g., Mower, 1960; Skinner, 1957) assumed that language learning would be subsumed under the general learning theory that behaviorists were striving to develop. Yet today the field has grown to such depth and complexity that a recent volume on the state of the art (Wanner & Gleitman, 1982) lists some 900 references, over half of them published in the last 10 years. The most that I can hope to do here is sketch some of the reasons for this phenomenal growth. What did Chomsky say that aroused such interest? What questions are researchers trying to answer?

Language development is a central issue in Chomsky's thought (e.g., 1965, 1972, 1980), bearing directly on the natural categories of the human mind. The issue arises from four assumptions. First, any grammar sufficient to generate the sentences of a natural language is a complex "system of many...rules of...different types organized in accordance with certain fixed principles of ordering and applicability and containing a certain fixed substructure" (1972, p. 75). Second, the descriptive predicates of this system (grammatical categories, phonological classes) are not commensurate with those of any other known system in the world or in the mind. Third, the data available to the child in the speech of others is "meager and degenerate." Fourth, no known theory of learning—least of all, a stimulus-response reinforcement theory of the kind scathingly criticized by Chomsky in his review (1959) of Skinner's *Verbal Behavior* (1957)—is adequate to account for a child's learning a language. Chomsky (1972) therefore
assigns to the mind an innate property, a schema constituting the "universal grammar" to which every language must conform. The schema is highly restrictive, so that the child's search for the grammar of the language it is learning will not be impossibly long.

Chomsky (1972) then divides the research task into three parts. First is the linguist's task: to define the essential properties of human language, the schema or universal grammar. Second is the psychologist's task of determining the minimal conditions that will trigger the child's innate linguistic mechanisms. The third task, closely related to the second, arises from the assumption that most of the utterances a child hears are not well formed. How then is the child to know which utterances to accept as evidence of the grammar it is searching for and which utterances to reject? The third task is therefore to discover the nature of the relation between a set of data and a potential grammar, sufficient to validate the grammar as a theory of the language being learned.

The proposition that language is an innate faculty of the human mind has a long history in Western thought from Plato to Darwin. The proposition is logically independent of any particular theory of language structure. Indeed, the entire enterprise of generative grammar might fail, yet leave the claim of innateness untouched. Certainly Chomsky's linguistic theories have been, and continue to be, a rich source of hypothesis and experiment in studies of language acquisition. However, his principle achievement in this area has been to force recognition that the learning of a language is an extraordinarily complex process with profound implications for the nature of mind. He has formulated the problem of language learning more precisely than ever before, spelling out its logical prerequisites in a fashion that promises to lead, given appropriate research, to a more precise specification of the innate "knowledge" that a child must bring to bear if it is ever to learn a language at all.

As we have noted, Chomsky's challenge precipitated a vast quantity of research. The first need was for data, for systematic descriptions of how language actually develops. Work initially concentrated on syntactic development (e.g., Brown, 1973), but in the past dozen years has expanded to include phonology (e.g., Yeni-Komshian, Kavanagh, & Ferguson, 1980), semantics (e.g., Carey, 1982; MacNamara, 1982) and pragmatics (e.g., Bates & MacWhinney, 1982). As data have accumulated, it has become possible to answer many questions and, of course, to ask many more.

When does language development begin? Can we isolate reliable stages of development across children? Do the same stages occur in different language environments? Is the input to the child truly "meager and degenerate"? Is the child really constructing a grammar? Is the process passive, or must the child actively engage itself? What is the role of imitation? Do we have to posit innate proclivities? If so, are they indeed purely linguistic? And so on.

To see the force of these questions, we must have a sense of the complexity of the task that faces a child learning its native language. From our discussion of the problems of speech perception and automatic speech recognition, it will be obvious that we have much to learn about how the infant discovers invariant phonetic and lexical segments in the speech signal. We still do not know how the infant learns the basic sound pattern of a language during its first two years of life and comes to speak its first few
dozen words. But let us set these puzzles aside and go straight to early syntax, where the bulk of child language research has been concentrated. The goal of this work has been to infer from a child's utterances (performance) what it "knows" (competence) about grammar, and the meanings encoded by grammar, at each stage of its development.

Consider, as an example, the sentence cited above, I want the apple we picked for supper, a sentence comfortably within the competence of a four-year-old child. What must a child know to produce such a sentence? We will look at three aspects of its structure to illustrate the basis of Chomsky's claim that grammatical categories do not map in any simple way onto the categories of general cognition.

(1) Word order. A child who utters the sentence evidently knows the standard subject-verb-object (SVO) order of English and so says, I want the apple. The child does not say as (transposing into English) a Turkish or Japanese child might say, I the apple want (SOV) or The apple I want (OSV). Presumably, the English-speaking child has long since learned that Adam loves Eve does not mean the same as Eve loves Adam. A Turkish or Japanese child, on the other hand, would have learned that uncertainties, due to variable word order, as to the underlying relations expressed in a sentence (who does what to whom) are resolved by attaching appropriate suffixes to subject and object (Slobin, 1982).

So far, the mapping between grammar and world, in the three languages, would seem to be arbitrary but direct. However, we are given pause by another phrase in our example, the apple we picked (=the apple that we picked). Here, in an object relative clause, the order of subject (we) and object (apple) is reversed, and the verb (picked) appears at the end, giving OSV. The switch from SVO (we picked that) to OSV (that we picked) is obligatory in English object relative clauses. Notice that, to apply this rule, a child cannot draw on any knowledge of the world; rather, it must (in some sense) know the grammatical structure of the sentence. We have here, then, another example of the structure dependence, noted above in our discussion of interrogatives.

(2) Use of the article. The child says, I want the apple, not I want an apple. Of course, if many apples had been picked, an apple would have been correct. The distinction between definite and indefinite articles seems natural to an English speaker. To a speaker of Russian, Chinese, or other languages in which articles are not used, the distinction might seem tiresome and unnecessary. In fact, rules for use of articles in English are complex and, with respect to the aspects of the world that they encode, seemingly arbitrary. Yet the rules are learned by the third or fourth year of life (Brown, 1973, p. 271).

(3) Noun phrases. As a final example, consider the noun phrase, the apple we picked. These four words (article + noun + adjectival phrase) form the grammatical object of the sentence. A child who utters them must already know the general rule for constructing noun phrases in English: the adjective goes before the noun (the red apple), not, as in French, after the noun (la pomme rouge). However, there is an exception to the rule: if the adjective is itself a phrase (that is, a relative clause: (that we picked), the adjective must follow the noun (the apple we picked, not the we picked apple). Once again, the child reveals in its utterance knowledge of a rule of English grammar that cannot be derived from knowledge of the world.
In short, there are solid grounds for believing that language structure (both at the level of sound pattern, or phonology, and at the level of syntax) may be sui generis. With this in mind, let us briefly review some of what we know about the course of development, with particular attention to the questions with which we began.

The infant is biologically prepared to distinguish speech from nonspeech at, or very soon after, birth. A double dissociation of the left cerebral hemisphere for perceiving speech and of the right hemisphere for perceiving nonspeech sounds within days of birth has been demonstrated both electrophysiologically (e.g., Molfeese, 1977) and behaviorally (e.g., Segalowitz & Chapman, 1980). Further, dozens of experiments in the past 10 years have shown that infants, in their first six months of life, can discriminate virtually any adult speech contrast from any language on which they are tested (e.g., [b] vs. [p], [d] vs. [g], [m] vs. [n], etc.) (Aslin, Pisoni, & Jusczyk, 1983; Eimas, 1982). There is also evidence that infants begin to recognize the function of such contrasts, to distinguish words in the surrounding language, during the second half of their first year (Werker, 1982). (For fuller review, see Studdert-Kennedy, 1986).

In terms of sound production, Oller (1980) has described a regular progression from simple phonation (0-1 months) through canonical babbling (7-10 months) to so-called variegated babbling (11-12 months). The phonetic inventory of babbled sounds is strikingly similar across many languages and even across hearing and deaf infants up to the end of the first year (Locke, 1983). These similarities argue for a universal, rather than language-specific, course of articularatory development.

However, around the end of the twelfth month, when the child produces its first words, the influence of the surrounding language becomes evident. From this point on, universals become increasingly difficult to discern, because whatever universals there may be are masked by surface diversity among languages. In this respect, the development of language differs from the development of, say, sensorimotor intelligence or mathematical ability (cf. Gelman & Brown, this volume). Nonetheless, we can already trace some regularities across children within a language and, to some lesser extent, across languages.

The most heavily studied stage of early syntactic development, in both English and some half-dozen other languages, is the so-called two-morpheme stage. Brown (1973) divides early development into five stages on the basis of mean length of utterance (MLU), measured in terms of the number of morphemes in an utterance. The stages are "not...true stages in Piaget's sense" (Brown, 1973, p. 58), but convenient, roughly equidistant points from MLU=2.00 through MLU=4.00. The measure provides an index of language development independent of a child's chronological age.

Of interest in the present context is that no purely grammatical description of Stage I (MLU=2.00, with an upper bound of 5.00) has been found satisfactory. Instead, the data are best described by a "rich interpretation," assigning a meaning or function to an utterance on the basis of the context in which it occurs. Brown lists 11 meanings for Stage I constructions, including: naming, recurrence (more cup), nonexistence (all gone egg), agent and action (Mommy go), agent and object (Daddy key), action and location (sit chair), entity and location (Baby table), possessor and possession (Daddy chair), entity and attribute (yellow bloo). Brown (1973)
proposes that these meanings "derive from sensorimotor intelligence, in Piaget's sense...[and] probably are universal in humankind but not...innate" (p. 201).

We should emphasize that these Stage I patterns reflect semantic, not grammatical, relations even though they may be necessary precursors to the grammatical relations that develop during Stage II (MLU=2.50, with an upper bound of 7.00). Brown (1973) traced the emergence of 14 grammatical morphemes in three Stage II English-speaking children. The morphemes included: prepositions (in, on), present progressive (I am playing), past regular (jumped), past irregular (broke), plural-s, possessive -s, third person -s (he jumps), and others. The remarkable finding was that all three children acquired the morphemes in roughly the same order (with rank order correlations between pairs of children of 0.86 or more). This result was confirmed in a study of 21 English-speaking children by de Villiers and de Villiers (1973).

However, unlike the meanings and functions of Stage I, the more or less invariant order of morpheme acquisition of Stage II has not been confirmed for languages other than English. Perhaps we should not expect that it will be. Languages differ, as we have seen, in the grammatical devices that they use to mark relations within a sentence. The devices used by one language to express a particular grammatical relation may be, in some uncertain sense, "easier" to learn than the devices used by another language for the same grammatical relation. Slobin (1982) has compared the ages at which four equivalent grammatical constructions are learned in Turkish, Italian, Serbo-Croatian, and English. In each case, the Turkish children developed more rapidly than the other children. If these results are valid and not mere sampling error, the "studies suggest that Turkish is close to an ideal language for early acquisition" (Slobin, 1982, p. 145).

Unless we suppose that Turkish parents are more attentive to their children's language than Italian, Serbo-Croatian, and English parents, we may take this result as further evidence that "selection pressures" (reinforcement) have little role to play in language learning. Brown and Hanlon (1970) showed some years ago that parents tend to correct the pronunciation and truth value, rather than the syntax, of their children's speech. Indeed, one of the puzzles of language development is why children improve at all. At each stage, the child's speech seems sufficient to satisfy its needs. Neither reinforcement nor imitation of adult speech suffices to explain the improvement. Early speech is replete with forms that the child has presumably never heard: two sheeps, we good, mine boot. These errors reflect not imitation, but over-generalization of rules for forming plurals, past tenses, and possessive adjectives.

We come then to a guiding assumption of much current research: Learning a first language entails active search for language-specific grammatical patterns (or rules) to express universal cognitive functions. The child may be helped in this by the relative "transparency" (Slobin, 1980) of the speech addressed to it--either because the language itself, like Turkish, is transparent and/or because adult speech to the child is conspicuously well formed. Several studies (e.g., Newport et al., 1977) have shown that the speech addressed to children tends not to be "degenerate." Yet the speech may be "meager" in the sense that relatively few instances suffice to trigger recognition of a pattern (Roeppe, 1982). Such rapid learning would seem to require a system specialized for discovering distinctive patterns of sound and syntax in any language to which a child is exposed.
Finally, it is worth remarking that all normal children do learn a language, just as they learn to walk. Western societies acknowledge this in their attitude to children who fail: We regard them as handicapped or defective, and we arrange clinics and therapeutic settings to help them. As Dale (1976) has remarked, we do not do the same for children who cannot learn to play the piano, do long division, or ride a bicycle. Of course, children vary in intelligence, but not until I.Q. drops below about 50 do language difficulties begin to appear (Lenneberg, 1967). Children at a given level of maturation also vary in how much they talk, what they talk about, and how many words they know. Where they vary little, it seems, is in their grasp of the basic principles of the language system—its sound structure and syntax.

Conclusion

The past 50 years have seen a vast increase in our knowledge of the biological foundations of language. Rather than attempt even a sampling of the issues raised by the research we have reviewed, let me end by emphasizing a point with which I began: the interplay between basic and applied research, and between research and theory.

The advances have come about partly through technological innovations, permitting, for example, physical analysis of the acoustic structure of speech and precise localization of brain abnormalities; partly through methodological gains in the experimental analysis of behavior; partly through growing social concern with the blind, the deaf, and otherwise language-handicapped persons. Yet these scattered elements would still be scattered had they not been brought together by a theoretical shift from description to explanation.

Perhaps the most striking aspect of the development is its unpredictability. Fifty years ago no one would have predicted that formal study of syntax would offer a theoretical framework for basic research in language acquisition, now a thriving area of modern experimental psychology, with important implications for treatment of the language-handicapped. No one would have predicted that applied research on reading machines for the blind would contribute to basic research in human phonetic capacity, lending experimental support to the formal linguistic claim of the independence of phonology and syntax. Nor, finally, would anyone have predicted that basic psycholinguistic research in American Sign Language would provide a unique approach to the understanding of brain organization for language and to testing the hypothesis, derived from linguistic theory, that language is a distinct faculty of the human mind.

Presumably, continued research in the areas we have reviewed and in related areas that we have not (such as the acquisition of reading, the motor control and coordination of articulatory action, second language learning), will consolidate our view of language as an autonomous system of nested subsystems (phonology, syntax). Beyond this lies the further task of unfolding the language system, tracing its evolutionary and ontogenetic origins in the nonlinguistic systems that surround it and from which, in the last analysis, it must derive. We would be rash to speculate on the diverse areas of research and theory that will contribute to this development.
Studdert-Kennedy: Some Developments in Research on Language Behavior

References


Müller, J. (1848). The physiology of the senses, voice, and muscular motion with the mental faculties (W. Baly, Trans.). New York: Walton and Maberly.


acquisition. Cambridge: Cambridge University Press.


