

The Specialization of the Language Hemisphere*

A. M. Liberman⁺

Haskins Laboratories, New Haven

The language hemisphere may be specialized to deal with grammatical recordings, which differ in important ways from other perceptual and cognitive processes. Their special function is to make linguistic information differentially appropriate for otherwise mismatched mechanisms of storage and transmission. At the level of speech we see the special nature of a grammatical code, the special model that rationalizes it, and the special mode in which it is perceived.

The fact that language is primarily on one side of the brain implies the question I will ask in this paper: how does language differ from the processes on the other side?¹ I will suggest, as a working hypothesis, that the difference is grammatical recoding, a conversion in which information is restructured, often radically, as it moves between the sounds of speech and the messages they convey. To develop that hypothesis, I will divide it into four more specific ones: grammatical codes have a special function; they restructure information in a special way; they are unlocked by a special key; and they are associated with a special mode of perception.

*Invited paper presented at the Intensive Study Program in the Neurosciences (Neurosciences Research Program, Massachusetts Institute of Technology) at Boulder, Colo., July 1972.

⁺Also University of Connecticut, Storrs, and Yale University, New Haven.

¹Language is the only cerebrally lateralized process I will be concerned with. I will not try to deal with its relation, if any, to other processes that may be in the same hemisphere, such as those underlying handedness or perception of fine temporal discriminations (Efron, 1965). Of course, we should understand cerebral specialization better if it could be shown that all the activities of one hemisphere were reflections of a single underlying design. (See, for example, Semmes, 1968.)

Acknowledgment: I am indebted to my colleagues at Haskins Laboratories, especially Franklin S. Cooper, Ignatius G. Mattingly, Donald Shankweiler, and Michael Studdert-Kennedy, for ideas, suggestions, and criticisms. Hans-Lukas Teuber, Brenda Milner, and Charles Liberman have also been very helpful. None of these people necessarily agrees with the views I express here.

[HASKINS LABORATORIES: Status Report on Speech Research SR-31/32 (1972)]

In talking about the function of grammatical codes, I will be concerned with language in general. Otherwise, I will limit my attention to speech and, even more narrowly, to speech perception. I do this partly because I know more about speech perception than about other aspects of language. But I am motivated, too, by the fact that more is known that bears on the purposes of this seminar. This becomes apparent when, in interpreting research on hemispheric specialization, we must separate processes that are truly linguistic from those that may only appear so. It becomes even more apparent when we try to frame experimental questions that might help us to discover, quite exactly, what the language hemisphere is specialized for. In any case, not so much is lost by this restriction of attention as might be supposed since, if recent arguments are accepted, speech perception is an integral and representative part of language, both functionally and formally (Liberman, 1970; Mattingly and Liberman, 1969).

THE SPECIAL FUNCTION OF GRAMMATICAL CODES: MAKING LINGUISTIC INFORMATION DIFFERENTIALLY APPROPRIATE FOR TRANSMISSION AND STORAGE

Perhaps the simplest way to appreciate the function of grammar is to consider what happens when we remember linguistic information. Should you try tomorrow to recall this lecture, we might expect, if what I say is sensible, that you would manage very well. But we can hardly conceive that you would reproduce exactly the strings of consonants and vowels, words, or sentences you will have heard. Nor can we suppose that your performance would be evaluated by any reasonable person in terms of the percentage of such elements you correctly recalled, or by the number of times your failure to recall lay merely in the substitution of a synonym for the originally uttered word. A judge of your recall would be concerned only with the extent to which you had captured the meaning of the lecture; he would expect a paraphrase, and that is what he would get.

Paraphrase is not a kind of forgetting but a normal part of remembering. It reflects the conversions that must occur if that which is communicated to us by language is to be well retained (and understood) or if that which we retain (and understand) is to be efficiently communicated. In the course of those conversions, linguistic information has at least three different shapes: an acoustic (or auditory) vehicle for transmission; a phonetic representation, consisting of consonants and vowels, appropriate for processing and storage in a short-term memory; and a semantic representation² that fits a nonlinguistic intellect and long-term memory. Of course, the conversions among these shapes would be of no special interest if they meant no more than the substitution of one unit for another--for example, a neural unit for an acoustic one--give or take the sharpening, distortions, and losses that must occur. But the facts of

²It may prove useful to make a distinction between a semantic representation, which presumably has linguistic structure, and some deeper base, which does not. We should suppose, then, that it is the less linguistic base that is stored in long-term memory, and that the semantic representation is synthesized from it. I believe, however, that such a distinction is not relevant to the purposes of this paper; moreover, there is no agreed-upon word to refer to the base form. I will, therefore, use "semantic representation" loosely to refer to whatever we might expect to find in long-term memory and the nonlinguistic intellect.

paraphrase imply far more than that kind of alphabetic encipherment. Since an accurate paraphrase need not, and usually does not, bear any physical resemblance to the originally presented acoustic (or auditory) signal, we must suppose that the information has been thoroughly restructured. It is as if the listener had stored a semantic representation that he synthesized or constructed out of the speech sounds, and then, on the occasion of recall, used the semantic representation as a base for synthesizing still another set of sounds. Plainly, these syntheses are not chaotic or arbitrary; they are, rather, constrained by rules of a kind that linguists call grammar. There is, therefore, a way to see the correspondence between the original and recalled information, or, indeed, between the transmitted and stored forms. But this can be done only by reference to the grammar, not by comparison of the physical properties of the two sets of acoustic events or of transforms performed directly on them. An observer who does not command the grammar cannot possibly judge the accuracy of the paraphrase.

Since my aim is to raise questions about the distinctiveness of language, I should pause here to ask whether paraphrase is unique. In visual memory, for example, is paraphrase even conceivable? Of course, the remembered scene one calls up in his mind's eye will usually differ from the original. But cannot the accuracy of recall always be judged by reference to the physical properties of the remembered scene, allowing, of course, for reversible transformations performed directly on the physical stimuli themselves? Except in the case of the most abstract art, about which there is notorious lack of agreement, can we ever say of two visual patterns that they correspond only in meaning, and, accordingly, that the correspondence between them can be judged only by reference to rules like those of grammar?

But I should return now to the function of grammatical recoding, which is the question before us. Why must the linguistic information be so thoroughly restructured if it is to be transmittable in the one case and storable in the other? The simple and possibly obvious answer is that the components for transmission and storage are grossly mismatched; consequently, they cannot deal with information in anything like the same form. I should suppose that the reason for the mismatch is that the several components developed separately in evolution and in connection with different biological activities. At the one end of the system is long-term memory, as well as the nonlinguistic aspects of meaning and thought. Surely, these must have existed before the development of language, much as they exist now in nonspeaking animals and, I dare say, in the nonlanguage hemisphere of man. At the other end of the system, the components most directly concerned with transmission--the ear and the vocal tract--had also reached a high state of development before they were incorporated as terminals in linguistic communication. [Important adaptations of the vocal tract did presumably occur in the evolution of speech, as has been shown (Lieberman, 1968, 1969; Lieberman and Cralin, 1971; Lieberman, Cralin, and Klatt, 1972; Lieberman, Klatt, and Wilson, 1969); however, these did not wholly correct the mismatch we are considering.] We might assume, then, following Mattingly (1972), that grammar developed as a special interface, joining into a single system the several components of transmission and intellect that were once quite separate. What are conceivably unique to language, to man, and to his language hemisphere are only grammatical codes. These are used to reshape semantic representations so as to make them appropriate, via a phonetic stage, for efficient transmission in acoustic form.

We should recognize, of course, that the consequences of being able to make those grammatical conversions might be immense, not merely because man can then more efficiently communicate his semantic representations to others, but also because he can, perhaps more easily than otherwise, move them around in his own head. If so, there may be thought processes that can be carried out only on information that has gone into the grammatical system, at least part way. We should also see that the nonlinguistic intellectual mechanisms might themselves have been altered in the course of evolutionary adaptations associated with the development of grammar. Indeed, exactly analogous adaptations did apparently take place at the other end of the system where, as has already been remarked, the vocal tract underwent structural changes that narrowed the gap between its repertory of shapes (and sounds) and that which was required by the nature of the phonetic representation at the next higher level. But such considerations do not alter my point, however much they may complicate it. We may reasonably suppose that the basic function of grammatical codes is to join previously independent components by making the best of what would otherwise be a bad fit.

At this point I should turn again to our question about the distinctiveness of language and ask whether the function of grammatical codes, as I described it here, is unique. Are there other biological systems in which different structures, having evolved independently, are married by a process that restructures the information passing between them? If not, then grammatical codes solve a biologically novel problem, and we should wonder whether it was in connection with such a solution that a new functional organization evolved in the left hemisphere.

But if we are to view grammar as an interface, we ought to see more clearly how bad is the fit that it corrects. For that purpose I will deal separately with two stages of the linguistic process: the interconversion between phonetic message and sound, which I will refer to throughout this paper as the "speech code," and then briefly with the part of language that lies between phonetic message and meaning.

The Phonetic Representation vs. the Ear and the Vocal Tract

At the phonetic level, language is conveyed by a small number of meaningless segments--roughly three dozen in English--called "phones" by linguists and well-known to us all as consonants and vowels. These phonetic segments are characteristic of all natural human languages and of no nonlinguistic communication systems, human or otherwise. Their role in language is an important one. When properly ordered, these few dozen segments convey the vastly greater number of semantic units; thus, they take a large step toward matching the demands of the semantic inventory to the possibilities of the vocal tract and the ear. They are important, too, because they appear to be peculiarly appropriate for storage and processing in short-term memory (Liberman, Mattingly, and Turvey, 1972). In the perception of speech the phonetic segments are retained in short-term memory and somehow organized into the larger units of words and phrases; these undergo treatment by syntactic and semantic processes, yielding up, if all goes well, something like the meaning the speaker intended. But if the larger organizations are to be achieved, the phonetic units must be collected at a reasonably high rate. (To see how important rate is, try to understand a sensible communication that is spelled to you slowly, letter by painful letter.) In fact, speaking speeds produce phonetic segments at rates of 8 to 20 segments per second, and research with artificially speeded speech (Orr, Friedman, and Williams, 1965) suggests that it is possible to perceive phonetic information at rates as high as 30 segments (that is, about seven words) per second.

Now if speech had developed from the beginning as a unitary system, we might suppose that the components would have been reasonably well matched. In that case there would have been no need for a radical restructuring of information--that is, no need for grammar--but only the fairly straightforward substitution of an acoustic segment for each phonetic one. Indeed, just that kind of substitution cipher has commonly been assumed to be an important characteristic of speech. But such a simple conversion would not work, in fact, because the requirements of phonetic communication are not directly met either by the ear or by the vocal tract.

Consider first the ear. If each phonetic unit were represented, as in an alphabet or cipher, by a unit of sound, the listener would have to identify from 8 to 30 segments per second. But such rates would surely strain, and probably overreach, the temporal resolving power of the ear. Consider next the requirement that the order of the segments be preserved. Of course, the listener would hardly be expected to order the segments if, at high rates, he could not even resolve them. We should note, however, that even at slower rates, and in cases where the identity of the sound segments is known, there is some evidence that the ear does not identify order well. Though this question has not been intensively investigated, data from the research of Warren, Obusek, Farmer, and Warren (1969) suggest that the requirements for ordering in phonetic communication would exceed the psychoacoustically determined ability of the ear by a factor of five or more.

Apparently, then, the system would not work well if the conversion from phonetic unit to sound were a simple one. We should suppose that this would be so for the reasons I just outlined. But the case need not rest on that supposition. In fact, there is a great deal of confirming evidence in the experience gained over many years through the attempts to develop and use acoustic (non-speech) alphabets. That experience has been in telegraphy--witness Morse code, which is a cipher or alphabet as I have been using the terms here--and much more comprehensively in connection with the early attempts to build reading machines for the blind. Even after considerable practice, users do poorly with those sound alphabets, attaining rates no better than one-tenth those which are achieved in speech (Coffey, 1963; Freiburger and Murphy, 1961; Nye, 1968; Studdert-Kennedy and Cooper, 1966).

Nor does the vocal tract appear to be better suited to the requirements of phonetic communication. If the sounds of speech are to be produced by movements of the articulatory organs, we should wonder where in the vocal tract we are going to find equipment for three dozen distinctive gestures. Moreover, we should wonder, since the order of the segments must be preserved, how a succession of these gestures can be produced at rates as high as one gesture every 50 msec.

The Phonetic vs. the Semantic Representations

Though appropriate for storage over the short term, the phonetic representation apparently does not fit the requirements of the long-term store or of the essentially nonlinguistic processes that may be associated with it. Those requirements are presumably better met by the semantic representation into which the phonetic segments are converted. Because of its inaccessibility, we do not know the shape of the information at the semantic level, which is a reason we do well, for our purposes, to concentrate our attention on the acoustic

and phonetic levels where we can more readily experiment. Still, some characteristics of the semantic representation can be guessed at. Thus, given the innumerable aspects of our experience and knowledge, we should suppose that the inventory of semantic units is very large, many thousands of times larger than the two or three dozen phonetic segments that transmit it. We should suppose, further, that however the semantic units may be organized, it is hardly conceivable that they are, like the phonetic segments, set down in ordered strings. At all events, the phonetic and semantic representations must be radically different, reflecting, presumably, the differences between the requirements of the processes associated with short- and long-term memory.

THE SPECIAL RESTRUCTURING PRODUCED BY THE SPEECH CODE:
SIMULTANEOUS TRANSMISSION OF INFORMATION ON THE SAME CUE

We can usefully think of grammatical coding as the restructuring of information that must occur if the mismatched components I have talked about are to work together as a single system. In developing that notion, I have so far spoken of three levels of linguistic information--semantic, phonetic, and acoustic--connected, as it were, by grammars that describe the relation between one level and the next. The phonetic and acoustic levels are linked by a grammar my colleagues and I have called the speech code. That is the grammar I shall be especially concerned with. But we should first place that grammar in the larger scheme of things, and establish some basis for demonstrating its resemblance to grammars of a more conventional kind. That has been done in some detail in recent reviews already referred to (Liberman, 1970; Mattingly and Liberman, 1969). Here I will offer the briefest possible account.

Exactly what we say about the more conventional grammars depends, of course, on which linguistic theory we choose. Fortunately, the choice is, for us, not crucial. Our purposes are well served by a very crude approximation to the transformational or generative grammar that is owing to Chomsky (1965). On that view, the conversion from semantic to phonetic levels is accomplished through two intermediate levels called "deep structure" and "surface structure." At each level--including also the phonetic, to which I have already referred--there are strings of segments (phones, words) organized into larger units (syllables, phrases). From one level to the next the organized information is restructured according to the rules of the appropriate grammar: syntax for the conversion from deep to surface, phonology for the conversion from surface to phonetic. It is not feasible to attempt an account of these grammars, even in broad terms. But I would point to one of the most general and important characteristics of the conversions they rationalize: between one level and the next there is no direct or easily determined correspondence in the number or order of the segments. Taking a simple example, we suppose that in the deep structure, the level closest to meaning, there are strings of abstract, word-like units which, when translated into the nearest kind of plain English, might say: The man is young. The man is tall. The man climbs the ladders. The ladders are shaky. According to the rules of syntax, and by taking advantage of referential identities, we should delete and rearrange the segments of the four deep sentences, emerging in due course at the surface with the single sentence: The tall young man climbs the shaky ladders. It is as if the first, second, and fourth of the deep sentences had been folded into the third, with the result that information about all four sentences is, at the surface, transmitted simultaneously and on the same words.

The information at the level of surface structure is in turn converted, often by an equally complex encoding, to the phonetic level. But I will only offer an example of one of the simplest aspects of that conversion which nevertheless shows that the information does change shape in its further descent toward the sounds of speech and also illustrates a kind of context-conditioned variation that grammatical conversions often entail. Consider in the word "ladders" the fate of the segment, spelled "s," that means "more than one." Its realization at the phonetic level depends on the segmental context: in our example, "ladders," it becomes [z]; in a word like "cats," it would be [s]; and in "house" it would be [əz].

The more obvious parts of grammar, and of the paraphrase which so strikingly reflects it, occur in the conversion between phonetic and semantic representations. But, as I have already suggested, there is another grammar, quite similar in function and in form, to be found in the speech code that connects the phonetic representation to sound. The characteristics of this code have been dealt with at some length in several recent papers (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967; Liberman, Mattingly, and Cooper, 1972). I will only briefly describe some of those characteristics now to show how they might mark speech perception and, by analogy, the rest of language as different from other processes.

How the Phonetic Message is Articulated: Matching the Requirements of Phonetic Communication to the Vocal Tract

Consider, again, that there are several times more segments than there are articulatory muscle systems capable of significantly affecting the vocal output. A solution is to divide each segment into features, so that a smaller number of features produces a larger number of segments, and then to assign each feature to a significant articulatory gesture. Thus, the phonetic segment [b] is uniquely characterized by four articulatory features: stop manner, i.e., rapid movement to or from complete closure of the buccal part of the vocal tract, which [b] shares with [d, g, p, t, k, m, n, ŋ] but not with other consonants; orality, i.e., closure of the velar passage to the nose, which [b] shares with [d, g, p, t, k], but not with [m, n, ŋ]; bilabial place of production, i.e., closure at the lips, which [b] shares with [p, m] but not with [d, g, t, k, n, ŋ]; and voiced condition of voicing, i.e., vocal fold vibration beginning simultaneously with buccal opening, which [b] shares with [d, g], but not with [p, t, k].

It remains, then, to produce these segments at high rates. For that purpose the segments are first organized into larger units of approximately syllabic size, with the restriction that gestures appropriate to features in successive segments be largely independent and therefore capable of being made at the same time or with a great deal of overlap. In producing the syllable, the speaker takes advantage of the possibilities for simultaneous or overlapping articulation, perhaps to the greatest extent possible. Thus, for a syllable like [bæg], for example, the speaker does not complete the lip movement appropriate for [b] before shaping the tongue for the vowel [æ] and then, only when that has been accomplished, move to a position appropriate for [g]. Rather, he overlaps the gestures, sometimes to such an extent that successive segments, or their component features, are produced simultaneously. In this way, co-articulation produces segments at rates faster than individual muscle

systems must change their states and is thus well designed, as Cooper (1966) has put it, to get fast action from relatively slow-moving machinery.

How the Co-Articulation of the Phonetic Message Produces the Peculiar Characteristics of the Speech Code

The grouping of the segments into syllables and the co-articulation of features represents an organization of the phonetic message, but not yet a very drastic encoding, since it is still possible to correlate isolable gestures with particular features. It is in the further conversions, from gestures to vocal-tract shapes to sounds, that the greater complications of the speech code are produced. For it is there that we find a very complex relation of gesture to vocal-tract shape and then, in the conversion from vocal-tract shape to sound, a reduction in the number of dimensions. The result is that the effects of several overlapped gestures are impressed on exactly the same parameter of the acoustic signal, thus producing the most important and complex characteristic of the speech code. That characteristic is illustrated in Figure 1, which is intended to demonstrate how several segments of the phonetic message are encoded into the same part of the sound. For that purpose, we begin with a simple syllable comprising the phonetic string [b] [æ] [g] and then, having shown its realization at the level of sound, we determine how the sound changes as we change the phonetic message, one segment at a time. The schematic spectrogram in the left-most position of the row at the top would, if converted to sound, produce an approximation to [bæg], which is our example. In that spectrogram the two most important formants--a formant is a concentration of acoustic energy representing a resonance of the vocal tract--are plotted as a function of time. Looking at only the second (i.e., higher) formant, so as to simplify our task, we try to locate the information about the vowel [æ]. One way to do that is to change the message from [bæg] to [bɔg] and compare the acoustic representations. The spectrogram for the new syllable [bɔg] is shown in the next position to the right, where, in order to make the comparison easier, the second formant of [bæg] is reproduced in dashed lines. Having in mind that [bæg] and [bɔg] differ only in their middle segments--that is, only in the vowels--we note that the difference between the acoustic signals is not limited, correspondingly, to their middle sections, but rather extends from the beginning of the acoustic signal to the end. We conclude, therefore, that the vowel information is everywhere in the second formant of the sound. To find the temporal extent of the [b] segment of our original syllable [bæg], we should ask, similarly, what the acoustic pattern would be if only the first segment of the phonetic message were now changed, as it would be, for example, in [gæg]. Looking, in the next position to the right, at that new syllable [gæg], we see that the change has produced a second-formant that differs from the original through approximately the first two-thirds of the temporal extent of the sound. A similar test for [g], the final consonant of our example, is developed at the right-hand end of the row; information about that segment exists in the sound over all of approximately the last two-thirds of its time course.

The general effect is illustrated in the single pattern in the lower half of the figure, which shows over what parts of the sound each of the three message segments extends. We see that there is no part of the sound that contains information about only one phonetic segment: at every point, the sound is carrying information simultaneously about at least two successive segments of the message, and there is a section in the middle where information is simultaneously

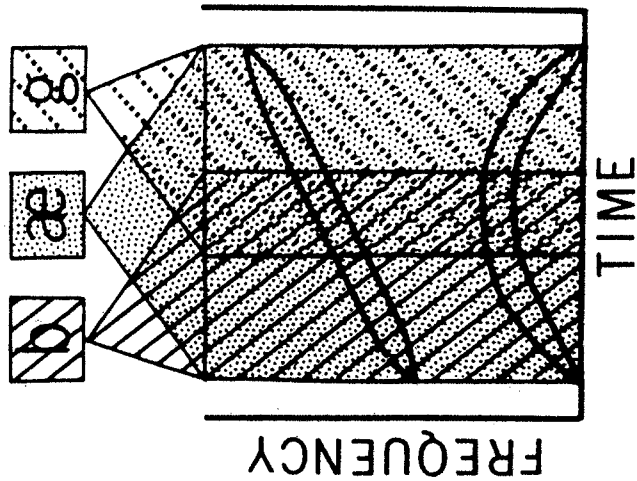
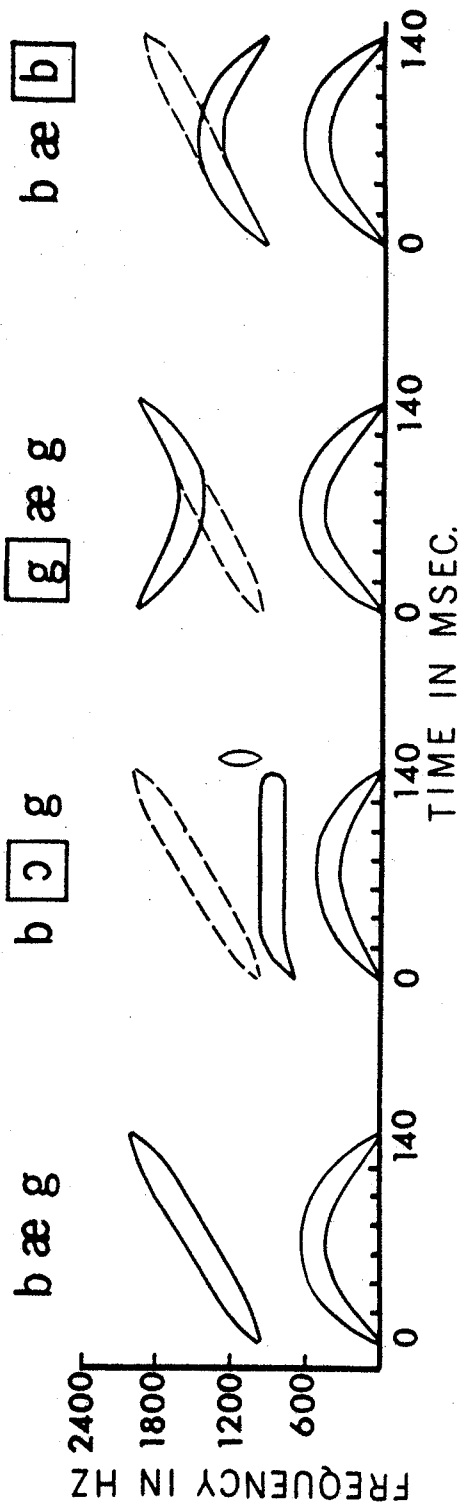


Fig. 1

Figure 1: Schematic spectrograms showing how the segments of the phonetic message are conveyed simultaneously on the same parameter of the sound.

available about all three. It is as if the initial and final consonants [b] and [g] had been folded into the vowel, much as the flanking deep-structure sentences of the earlier syntactic example were folded into that middle sentence that served, like the vowel, as a core or carrier.

Given that information about successive segments of the message is often carried simultaneously on the same parameter of the signal, the acoustic shape of a cue for a particular segment (or feature) will necessarily be different in different contexts. To see that this is so we should look again at the figure, but instead of noting, as we did before, that changing only the middle segment caused a change in the entire acoustic signal, we should see now that, though we retained two of the three original message segments, we nevertheless left no part of the acoustic signal intact. That is to say that the acoustic cues for [b] and [g] are very different in the contexts of the different vowels into which they are encoded. Such context-conditioned variation, similar perhaps to that we noted in the phonology, is often very great, not only for a consonant segment with different vowels, as in the example offered here, but also for different positions in the syllable, different kinds of syllable boundaries, different conditions of stress, and so on. (See, for example, Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967.)

Thus, as in the conversions between other levels of the language structure, the connection between phonetic message and sound is that of a very complex code, not an alphabet or substitution cipher. Information about successive segments of the message is often encoded into a single acoustic event with the result that there is no direct or easily calculated correspondence in segmentation, and the resulting variation in the shape of the acoustic cue can be extreme. At the levels of phonetic and acoustic representations, those characteristics define what I mean by a grammatical code.

But is the speech code--and, by extension, the other grammatical codes--unique? In visual and auditory perception, the relations between stimulus pattern and perceived response may be just as complex as those of speech, but they appear, as a class, to be different. I find it difficult to characterize the difference in general terms beyond saying that, apart from speech perception, we do not find the kind of simultaneous transmission that requires the perceiver to process a unitary physical event so as to recover the two or three discrete perceptual events that are encoded in it.

How the Speech Code Matches the Requirements of Phonetic Communication to the Properties of the Ear

I remarked earlier that we can and do hear speech at rates that would appear to overreach the resolving power of the ear if each phone were transmitted by a unit sound. But we have seen that the phones are not transmitted in that direct way; they are, rather, converted so as to encode several phones into the same acoustic unit. Though this produces a great complication in the relation between signal and message, and one that will have to be dealt with by a correspondingly complex decoder, it serves the important purpose of reducing significantly the number of discrete acoustic events that must be heard, and thus makes it possible to perceive phonetic information at reasonable rates. Given that the segments are encoded into units of approximately syllabic size, we should suppose that the limit on perception is set, not by the number of phonetic segments per unit time, but more nearly by the number of syllables.

I also remarked earlier on another way in which the ear appears to be ill suited to the requirements of phonetic communication: a listener must identify the order of phonetic segments, yet in ordinary auditory perception he cannot do that well. The solution to this problem that is offered by the speech code is that order is often marked, not only by time of occurrence, but also by context-conditioned variations in the shape of the cue. Thus, because of the kind of encoding that occurs, a primary acoustic cue for the two b's in [bæb] will be mirror images of each other. In words like [tæks] and [tæsk] the acoustic cues for [k] will have very different shapes, again because of co-articulation. Hence, the speech code offers the listener the possibility of constructing (or, more exactly, reconstructing) the order of the segments out of information which is not simply, or even primarily, temporal.

More and Less Encoded Aspects of Speech

An important characteristic of the speech code, especially in relation to questions about hemispheric specialization, is that not all parts of the speech signal bear a highly encoded relation to the phonetic message. In slow to moderate articulation, vowels and fricatives, for example, are sometimes represented by a simple acoustic alphabet or cipher: there are isolable segments in which information about only one phonetic segment is carried, and there may be little variation in the shape of the acoustic cues with changes in context. Segments belonging to the classes liquids and semivowels can be said to be grammatically encoded to an intermediate degree. Though these segments cannot be isolated in the speech signal (except for r-colored vowels), they do have brief steady-state portions, even in rapid articulation.

Nongrammatical Complications in the Relation between Acoustic and Phonetic Levels

There are several characteristics of speech apart from its encodedness that might require special treatment in perception. One is that the speech signal seems very poorly designed, at least from an engineering point of view. The acoustic energy is not concentrated in the information-bearing parts of the sound but is, rather, spread quite broadly across the spectrum. Moreover, the essential acoustic cues are, from a physical point of view, among the most indeterminate. Thus, the formant transitions, so important in the perception of most consonants, are rapid changes in the frequency position of a resonance which, by their nature, scatter energy.

Another kind of difficulty arises from the gross variations in vocal tract dimensions among men, women, and children. A consequence is that the absolute values of the formant cues will be different depending on the sex and size of the speaker. Obviously, some kind of calibration is necessary if the listener is to perceive speech properly.

What, Then, Is the Language Hemisphere Specialized for?

I have suggested, as a working hypothesis, that the distinctive characteristic of language is not meaning, thought, communication, or vocalization, but, more specifically, a grammatical recoding that reshapes linguistic information so as to fit it to the several originally nonlinguistic components of the system. That hypothesis may be useful in research on hemispheric specialization for language because it tells how we might make the necessary distinction

between that which is linguistic and that which is not. Our aim, then, is to discover whether it is, in fact, the processes of grammatical recoding that the language hemisphere is specialized for. That will be hard to do at the level closest to the semantic representation because we cannot, at that end of the language system, so easily define the boundary between grammatical coding and the presumably nonlinguistic processes it serves. But in speech, and especially in speech perception, we can be quite explicit. As a result, we can ask pointed questions and, because appropriate techniques are available, get useful answers. I will offer a few examples of such questions and answers. Because the experiments I will talk about represent a large and rapidly growing class, I should emphasize that, for the special purposes of this paper, I will describe only a few.

Speech vs. nonspeech. After investigations of people with cortical lesions, including especially the studies by Milner (1954, 1958), had indicated that perception of speech and nonspeech might be primarily on opposite sides of the head, Kimura (1961a, 1961b) pioneered the development of an experimental technique that permits us to probe this possibility with normal people. Adapting for her purposes a method that had been used earlier by Broadbent (1956), Kimura presented spoken digits dichotically, one to one ear and a different one to the other ear. She discovered that most listeners heard better the digits presented to the right ear. It was subsequently found, by her and others, that the same effect is obtained with nonsense syllables, including those that differ in only one phonetic segment or feature (Kimura, 1967; Shankweiler and Studdert-Kennedy, 1967). When the stimuli are musical melodies or complex nonspeech sounds, the opposite effect, a left-ear advantage, is obtained (Kimura, 1964). On the assumption that the contralateral auditory representation is stronger than the ipsilateral, especially under conditions of dichotic competition, Kimura interpreted these findings to reflect left-hemisphere processing of the speech signals and right-hemisphere processing of the others. In any case, many studies now support the conclusion that the ear advantages are reliable reflections of the functional asymmetry of the cerebral hemispheres. (For summaries see: Kimura, 1967; Shankweiler, 1971; Studdert-Kennedy and Shankweiler, 1970.)

Auditory vs. phonetic processing. If, as seems reasonable, the right-ear advantage for speech is interpreted to reflect the work of some special device in the left hemisphere, we should ask whether that device is specialized for grammatical decoding or for something else. Consider, then, a case such as the stop consonants. As I pointed out earlier, these phonetic segments are encoded grammatically in the exact sense that there is no part of the acoustic signal that carries information only about the consonant; the formant transitions, which contain all the information about the consonant, are simultaneously providing information about the following vowel. Any device that would perceive the segments correctly must deal with that grammatical code. Conceivably, that is what the device in the language hemisphere is specialized for. But there are other, nongrammatical jobs to be done and, accordingly, other possibilities. Among these are the tasks I referred to earlier when I spoke of the need to clean up the badly smeared speech signal, to track the very rapid frequency modulations (formant transitions) that are such important cues, and to calibrate for differences in vocal-tract size. Though not grammatical according to our definition, these tasks confront the listener only in connection with speech. They might, therefore, be more closely associated with the language hemisphere than those other auditory processes that must underlie the perception of all sounds, speech and nonspeech alike. But that is precisely the kind of issue that can be settled experimentally.

Several investigators (for example Darwin, 1971; Haggard, 1971; Shankweiler and Studdert-Kennedy, 1967; Studdert-Kennedy and Shankweiler, 1970; Studdert-Kennedy, Shankweiler, and Pisoni, 1972) have suggested and considerably refined questions like those I posed in the preceding paragraph and, in a number of ingenious experiments, found answers. I cannot here describe, or even summarize, these generally complex studies except to say that they provide some support for the notion that in speech perception the language hemisphere extracts phonetic features, which is to say in the terminology of this paper that it does grammatical decoding. There is, however, an experiment by Darwin (1971) which suggests that the language hemisphere may also be responsible for normalizing the acoustic signal to take account of the complications produced by the differences among speakers in length of the vocal tract. That finding indicates that our hypothesis is, at best, incomplete. Of course, we can hope to discover a mechanism general enough to include both vocal-tract normalization and grammatical decoding, the more so since these processes are so intimately associated with each other and with nothing else. Meanwhile, we can proceed to find out by experiment whether the language hemisphere is responsible for the other nongrammatical tasks, however closely or remotely they may be associated with speech. Perhaps an example of such an experiment will clarify the question and also our hypothesis.

Imagine a set of stop-vowel syllables [ba, da, ga] synthesized in such a way that the only distinguishing acoustic cue is the direction and extent of the first 50 msec of the second-formant transition, the rapid frequency modulation referred to earlier. Suppose, now, that we present these dichotically--that is, [ba], for example, to one ear, [da] to the other--in randomly arranged pairs and, as usual, get the right-ear advantage that is presumed to reflect left-hemisphere processing. On the hypothesis proposed here, we should say that these signals were being processed in the language hemisphere because they required grammatical decoding. In that case, we should have in the language hemisphere a device that is quite properly part of a linguistic system. There is, however, an alternative, as I have already implied, which is that the language hemisphere is specialized not for grammatical decoding but for responding to a particular class of auditory events, specifically the rapid frequency modulations of the second-formant transitions that are, in the stimuli of the experiment, the only acoustic cues. In that case, the left hemisphere would be said, at least in this respect, to be specialized for an auditory task, not a linguistic one. An experiment that helps to decide between these possibilities would go as follows. First, we remove from the synthetic syllables the second-formant transition cues and present them in isolation. When we do that we hear, not speech, but more or less distinguishable pitch glides or bird-like chirps. Now, given that there is a right-ear (left-hemisphere) advantage when these formant-transition cues are in a speech pattern, we determine the ear advantage when they are presented alone and not heard as speech. Donald Shankweiler, Ann Syrdal, and I (personal communication) have been doing that experiment. The results so far obtained are not wholly convincing, because, owing largely to the difficulty our listeners have in identifying the transition cues alone, the data are quite noisy. So far as the results can be interpreted, however, they suggest that the second-formant transitions in isolation produce a left-ear advantage, in contrast to the right-ear advantage obtained when those same transitions cued the perceived distinctions along [ba, da, ga]. If that result proves reliable, we should infer that the language hemisphere is specialized for a linguistic task of grammatical decoding, not for the auditory task of tracking formant transitions.

A clearer answer to essentially the same question, arrived at by a very different technique, is to be found in a recent doctoral dissertation by Wood (in preparation). He first replicated an earlier study (Wood, Goff, and Day, 1971) in which it had been found that evoked potentials were exactly the same in the right hemisphere whether the listener was distinguishing two syllables that differed only in a linguistically irrelevant dimension (in this case [ba] on a low pitch vs. [ba] on a high pitch) or in their phonetic identity ([ba] vs. [da] on the same pitch) but the evoked potentials in the left hemisphere were different in the two cases. From that result it had been inferred that the processing of speech required a stage beyond the processing of the nonlinguistic pitch parameter and, more important, that the stage of speech processing occurred in the left hemisphere. Now, in his dissertation, Wood has added several other conditions. Of particular interest here is one in which he measured the evoked potentials for the isolated acoustic cues, which were, as in the experiment described above, the second-formant transitions. The finding was that the isolated cue behaved just like the linguistically irrelevant pitch, not like speech. This suggests, as does the result we have so far obtained in the analogous dichotic experiment, that the processor in the language hemisphere is specialized, not for a particular class of auditory events, but for the grammatical task of decoding the auditory information so as to discover the phonetic features.

More vs. less encoded elements. As we saw earlier, only some phonetic segments--for example, [b, d, g]--are always grammatically encoded in the sense that information about them is merged at the acoustic level with information about adjacent segments. Others, such as the fricatives and the vowels, can be, and sometimes are, represented in the sound as if in a substitution cipher; that is, pieces of sounds can be isolated which carry information only about those segments. Still others, the liquids and semi-vowels, appear to have an intermediate degree of encodedness. We might suppose that only the grammatically encoded segments need to be processed by the special phonetic decoder in the left hemisphere; the others might be dealt with adequately by the auditory system. It is of special interest, then, to note the evidence from several studies that the occurrence or magnitude of the right-ear advantage does depend on encodedness (Darwin, 1971; Haggard, 1971; Shankweiler and Studdert-Kennedy, 1967). Perhaps the most telling of these experiments is a very recent one by Cutting (1972). He presented stop-liquid-vowel syllables dichotically--e.g., [kre] to one ear, [glæ] to the other--and found for the stops that almost all of his subjects had a right-ear advantage, while for the vowels the ear advantage was almost equally divided, half to the right ear and half to the left; the results with the liquids were intermediate between those extremes.

We might conclude, again tentatively, that the highly encoded aspects of speech--those aspects most in need of grammatical decoding--are always (or almost always) processed in the language hemisphere. The unencoded or less highly encoded segments may or may not be processed there. We might suppose, moreover, that some people tend to process all elements of language linguistically while others use nonlinguistic strategies wherever possible. If that is so, could it account for at least some of the individual differences in "degree" of ear advantage that turn up in almost all investigations?

Primary vs. secondary speech codes; cross codes. People ordinarily deal with the complications of the speech code without conscious awareness. But awareness of some aspects of speech, such as its phonetic structure, is sometimes

achieved. When that happens secondary codes can be created, an important example being language in its alphabetically written form. This written, secondary code is not so natural as the primary code of speech, but neither is it wholly unnatural, since it presumably makes contact with a linguistic physiology that is readily accessible when reading and writing are acquired. Research suggests that the contact is often (if not always) made at the phonetic level (Conrad, 1972); that is, that which is read is recoded into a (central) phonetic representation. If so, then we might expect to see the consequences in studies of hemispheric specialization for the perception of written language, as indeed we do (Milner, 1967; Umiltà, Frost, and Hyman, 1972).

In addition to the complications of secondary codes, there are special problems arising out of the tendency, under some conditions, to cross-code non-linguistic experience into linguistic form. As found in a recent experiment by Conrad (1972), for example, confusions in short-term memory for pictures of objects were primarily phonetic, not visual (or optical). Such results do not reveal the balance of nonlinguistic and linguistic processes, but they make it nonetheless evident that in the perception of pictures and, perhaps, of other kinds of nameable patterns, too, some aspects of the processing might be linguistic and therefore found in the language hemisphere.

A SPECIAL KEY TO THE CODE: THE GRAMMAR OF SPEECH

If the speech code were arbitrary--that is, if there were no way to make sense of the relation between signal and message--then perception could only be done by matching against stored templates. In that case there could be no very fundamental difference between speech and nonspeech, only different sets of templates. Of course, the number of templates for the perception of phonetic segments would have to be very large. It would, at the least, be larger than the number of phones because of the gross variations in acoustic shape produced by the encoding of phonetic segments in the sound; but it would also be larger than the number of syllables, because the effects of the encoding often extend across syllable boundaries and because the acoustic shape of the syllable varies with such conditions as rate of speaking and linguistic stress.

But grammatical codes are not arbitrary. There are rules--linguists call them grammars--that rationalize them. Thus, in terms of the Chomsky-like scheme I sketched earlier, the grammar of syntax tells us how we can, by rule, reshape the string of segments at the level of deep structure so as to arrive at the often very different string at the surface. In the case of the speech code we have already seen the general outlines of the grammatical key: a model of the articulatory processes by which the peculiar but entirely lawful complications of the speech code come about. The chief characteristic and greatest complication of the speech code, it will be recalled, is that information about successive segments of the message is carried simultaneously on the same acoustic parameter. To rationalize that characteristic we must understand how it is produced by the co-articulation I described earlier. Though crude and oversimple, that account of co-articulation may nevertheless have shown that a proper model of the process would explain how the phonetic message is encoded in the sound. Such a proper model would be a grammar, the grammar of speech in this case. It would differ from other grammars--for example, those of syntax and phonology--in that the grammar of speech would be a grammar done in flesh and blood, not, as in the case of syntax, a kind of algebra with no describable physiological correlates. Because the grammar of speech would correspond to an actual process,

it is tempting to suppose that the understanding of the speech code it provides is important, not just to the inquiring scientist, but also to the ordinary listener who might somehow use it to decode the complex speech sounds he hears. To yield to that temptation is to adopt what has been called a "motor theory of speech perception," and then to wonder if the language hemisphere is specialized to provide a model of the articulatory processes in terms of which the decoding calculations can be carried out.

One finds more nearly direct evidence for a motor theory when he asks which aspect of speech, articulatory movement or sound, is more closely related to its perception. That question is more sensible than might at first appear because the relation between articulation and sound can be complex in the extreme. Thus, as I have already indicated, the section of sound that carries information about a consonant is often grossly altered in different vowel contexts, though the consonant part of the articulatory gesture is not changed in any essential way. Following articulation rather than sound, the perception in all these cases is also unchanged (Liberman, 1957; Liberman, Delattre, and Cooper, 1952; Lisker, Cooper, and Liberman, 1962). Though such findings support a motor theory, I should note that only a weak form of the theory may be necessary to account for them. That is, they may only suggest that the perception of phonetic features converges, in the end, on the same neural units that normally command their articulation; in that case the rest of the processes underlying speech perception and production could be quite separate.

Evidence of a different kind can be seen in the results of a recent unpublished study by L. Taylor, B. Milner, and C. Darwin. Testing patients with excisions of the face area in the sensori-motor cortex of the left hemisphere, these investigators found severe impairments in the patients' ability to identify stop consonants (by pointing to the appropriate letter printed on a card) in nonsense-syllable contexts, though the pure-tone audiograms and performance on many other verbal tasks were normal. Patients with corresponding damage in the right hemisphere, and those with temporal or frontal damage in either hemisphere, were found not to differ from normal control subjects. It is at least interesting from the standpoint of a motor theory that lesions in the central face area did produce an inability to identify encoded stop consonants, though, as the investigators have pointed out, the exact nature of the impairment, whether of perception or of short-term memory, will be known only after further research.

The idea that the left hemisphere may be organized appropriately for motor control of articulation is in the theory of hemispheric specialization proposed by Semmes (1968). It is, perhaps, not inconsistent with her theory to suppose, as I have, that the organization of the language hemisphere makes a motor model more available to perceptual processes. But one might, on her view, more simply assume that lateralization for language arose primarily for reasons of motor control. This would fit with the suggestion by Levy (1969) that, to avoid conflict, it would be well not to have bilaterally issued commands for unilateral articulations. In that respect speech may be unique, as Evarts (personal communication) has pointed out, since other systems of coordinated movements ordinarily require different commands to corresponding muscles on the two sides. Conceivably, then, motor control of speech arose in one hemisphere in connection with special requirements like those just considered, and then everything else having to do with language followed. This assumption has the virtue of simplicity, at least in explaining how language got into one hemisphere in the first place.

Moreover, it is in keeping with a conclusion that seems to emerge from the research on patients with "split" brains, which is that, of all language functions, motor control of speech is perhaps, most thoroughly lateralized (Sperry and Gazzaniga, 1967).

At all events, though the grammar of speech makes sense of the complexly encoded relation between phonetic message and sound, it does not tell us how the decoding might be carried out. Like the other grammars of phonology and syntax, the grammar of speech works in one direction only, downward; the rules that take us from phonetic message to sound do not work in reverse. Indeed, we now know the downward-going rules well enough, at least in acoustic form, to be able to use them (via a computer) to generate intelligible speech automatically from an input of (typed) phonetic segments (Mattingly, 1968, 1971). But we do not know how to go automatically in the reverse direction, from speech sounds to phonetic message, except perhaps via the roundabout route of analysis-by-synthesis--that is, by guessing at the message, generating (by rule) the appropriate sound, and then testing for match (Stevens, 1960; Stevens and Halle, 1967).

Still, I should think that we decode speech with the aid of a model that is, in some important sense, articulatory. If so, we might suppose that the functional organization of the left hemisphere is peculiarly appropriate for the conjoining of sensory and motor processes that such a model implies.

Having said that the speech code is rationalized by a production model, I should ask whether in this respect it differs from the relations between stimulus and perception in other perceptual modalities. I think perhaps it does. In visual and auditory perception of nonverbal material the complex relations between stimulus and perception are also "ruly" rather than arbitrary, but the rules are different from those of the speech code if only because the complications between stimulus and perception in the nonspeech case do not come about as a result of the way the human perceivers produce the stimulus: the very great complications of shape constancy, for example, are rationalized, not in terms of how a perceiver makes those shapes, but by the rules of projective geometry. This is not to say that motor considerations are unimportant in nonspeech perception. Obviously, we must, in visual perception, take account of head and eye movements, else the world would appear to move when it should stand still (Teuber, 1960:1647-1648). But in those cases the motor components must be entered only as additional data to be used in arriving at the perception; the perceptual calculations themselves would be done in other terms.

I wonder, too, if the fact that the speech rules work in only one direction makes them different from those that govern other kinds of perception. In the case of shape constancy, for example, we know that one can, by the rules of geometry, calculate the image shape on the retina if he knows the shape of the stimulus object and its orientation. That would be analogous to using the grammar of speech to determine the nature of the sound, given the phonetic message. But in shape constancy, it would appear that the calculations could be made in reverse--that is, in the direction of perception. Knowing the image shape on the retina and the cues for orientation, one ought to be able to calculate directly the shape of the object. If so, then there would be no need in shape constancy, and conceivably in other kinds of nonspeech perception, for a resort to analysis-by-synthesis if the perceptual operations are to be done by calculation; rather, the calculations could be performed directly.

SPECIAL CHARACTERISTICS OF PERCEPTION IN THE SPEECH MODE

A commonplace observation about language is that it is abstract and categorical. That means, among other things, that language does not fit in any straightforward or isomorphic way onto the world it talks about. We do not use longer words for longer objects, or, less appropriately, louder words for bluer objects. If we change only one phonetic segment out of four in a word, we do not thereby create a word less different in meaning than if we had changed all four. Apart from onomatopoeia and phonetic symbolism, which are among the smallest and least typical parts of language, we do not use continuous linguistic variations to represent the continuous variations of the outside world.

It is of interest, then, to note that in the case of the encoded phonetic segments speech perception, too, is abstract. In listening to the syllable [ba], for example, one hears the stop consonant as an abstract linguistic event, quite removed from the acoustic and auditory variations that underlie it. He cannot tell that the difference between [ba] and [ga] in simplified synthetic patterns is only a rising frequency sweep in the second formant of [b] compared with a falling frequency sweep in the second formant of [g]. But if those frequency sweeps are removed from the syllable context and sounded alone, they are heard as rising and falling pitches, or as differently pitched "chirps," just as our knowledge of auditory psychophysics would lead us to expect. Perception in that auditory mode follows the stimulus in a fairly direct way; in that sense, and in contrast to the perception of speech, it is not abstract.

Perception of the encoded segments of speech is, as a corollary of its abstractness, also categorical. Thus, if we vary a sufficient acoustic cue for [b, d, g] in equal steps along a physical continuum, the listener does not hear step-wise changes but more nearly quantal jumps from one perceived category to another. This categorical perception has been measured by a variety of techniques and has been given several different but not wholly unrelated interpretations (Conway and Haggard, 1971; Fry, Abramson, Eimas, and Liberman, 1962; Fujisaki and Kawashima, 1969; Liberman, Harris, Hoffman, and Griffith, 1957; Pisoni, 1971; Stevens, Liberman, Ohman, and Studdert-Kennedy, 1969; Vinegrad, 1970). It characterizes the grammatically encoded segments (e.g., stop consonants), as I have indicated, but not the segments (e.g., the vowels in slow articulation) that are, as I noted earlier, represented in the acoustic signal as if by an alphabet or substitution cipher. Moreover, categorical perception cannot be said to be characteristic of a class of acoustic (and corresponding auditory) events, because the acoustic cues are perceived categorically only when they cue the distinctions among speech sounds; when presented in isolation and heard as nonspeech, their perception is more nearly continuous (Mattingly, Liberman, Syrdal, and Halwes, 1971).

At all events, the grammatically encoded aspects of speech do appear to be perceived in a special mode. That mode is, like the rest of language, abstract, categorical, and, perhaps more generally, nonrepresentational. Does this not present a considerable contrast to nonverbal visual and auditory perception? For all the abstracting that special detector mechanisms may do in vision or hearing, perception in those modes seems nevertheless to be more nearly isomorphic with the physical reality that occasions it. If that is truly a difference between grammatical and nongrammatical perception, it may be yet another reflection of the different organizations of the cerebral hemispheres.

SUMMARY

The aim of this paper is to suggest that the language hemisphere may be specialized to deal with grammatical coding, a conversion of information that distinguishes language from other perceptual and cognitive processes. Grammatical coding is unique, first, in terms of its function, which is to restructure information so as to make it appropriate for long-term storage and (nonlinguistic) cognitive processing at the one end of the system and for transmission via the vocal tract and the ear at the other.

To see further how grammatical restructurings are unique, we should look, more narrowly, at the speech code, the connection between phonetic message and sound. There we see a grammatical conversion that produces a special relation between acoustic stimulus and perception: information about successive segments of the perceived phonetic message is transmitted simultaneously on the same parameter of the sound. On that basis we can tentatively distinguish that which is grammatical or linguistic from that which is not. Then, by taking advantage of recently developed experimental techniques, we can discover to what extent our hypothesis about hemispheric specialization is correct and how it needs to be modified.

The speech code is unique in still other ways that may be correlates of the special processes of the language hemisphere. Thus, the speech code requires a special key. To understand the relation between acoustic stimulus and perceived phonetic message, one must take account of the manner in which the sound was produced. Conceivably, the language hemisphere is specialized to provide that "understanding" by making available to the listener the appropriate articulatory model.

The speech code is unique, too, in that it is associated with a special mode of perception. In that mode perception is categorical, digital, and most generally, nonrepresentational. Perhaps these perceptual properties reflect the specialized processes of the language hemisphere.

REFERENCES

- Broadbent, D. E. (1956) Successive responses to simultaneous stimuli. *Quart. J. Exp. Psychol.* 8, 145-162.
- Chomsky, N. (1965) Aspects of the Theory of Syntax. (Cambridge, Mass.: MIT Press).
- Coffey, J. L. (1963) The development and evaluation of the Batelle aural reading device. Proceedings of the International Congress on Technology and Blindness I. (New York: American Foundation for the Blind) 343-360.
- Conrad, R. (1972) Speech and reading. In Language by Ear and by Eye, ed. by J. F. Kavanagh and I. G. Mattingly. (Cambridge, Mass.: MIT Press) 205-240.
- Conway, D. A. and M. P. Haggard. (1971) New demonstrations of categorical perception. In Speech Synthesis and Perception Progress Report No. 5. (Cambridge, England: Psychological Laboratory) 51-73.
- Cooper, F. S. (1966) Describing the speech process in motor command terms. *J. Acoust. Soc. Amer.* 39, 1121-A. (Also in Haskins Laboratories Status Report on Speech Research SR-5/6, 2.1-2.27.)
- Cutting, J. E. (1972) A parallel between encodedness and the magnitude of the right ear effect. Haskins Laboratories Status Report on Speech Research SR-29/30, 61-68.

- Darwin, C. J. (1971) Ear differences in recall of fricatives and vowels. *Quart. J. Exp. Psychol.* 23, 46-62.
- Efron, R. (1965) The effect of handedness on the perception of simultaneity and temporal order. *Brain* 86, 261-284.
- Freiberger, J. and E. G. Murphy. (1961) Reading machines for the blind. IRE Professional Group on Human Factors in Electronics HFE-2, 8-19.
- Fry, D. B., A. S. Abramson, P. D. Eimas, and A. M. Liberman. (1962) The identification and discrimination of synthetic vowels. *Lang. Speech* 5, 171-189.
- Fujisaki, H. and T. Kawashima. (1969) On the modes and mechanisms of speech perception. In Annual Report No. 1. (Tokyo: University of Tokyo, Division of Electrical Engineering, Engineering Research Institute) 67-73.
- Haggard, M. P. (1971) Encoding and the REA for speech signals. *Quart. J. Exp. Psychol.* 23, 34-45.
- Kimura, D. (1961a) Some effects of temporal-lobe damage on auditory perception. *Canad. J. Psychol.* 15, 156-165.
- Kimura, D. (1961b) Cerebral dominance and perception of verbal stimuli. *Canad. J. Psychol.* 15, 166-171.
- Kimura, D. (1964) Left-right differences in the perception of melodies. *Quart. J. Exp. Psychol.* 16, 355-358.
- Kimura, D. (1967) Functional asymmetry of the brain in dichotic listening. *Cortex* 3, 163-178.
- Levy, J. (1969) Possible basis for the evolution of lateral specialization of the human brain. *Nature* 224, 614-615.
- Liberman, A. M. (1957) Some results of research on speech perception. *J. Acoust. Soc. Amer.* 29, 117-123.
- Liberman, A. M. (1970) The grammars of speech and language. *Cog. Psychol.* 1, 301-323.
- Liberman, A. M., F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy. (1967) Perception of the speech code. *Psychol. Rev.* 74, 431-461.
- Liberman, A. M., P. C. Delattre, and F. S. Cooper. (1952) The role of selected stimulus variables in the perception of the unvoiced stop consonants. *Amer. J. Psychol.* 65, 497-516.
- Liberman, A. M., K. S. Harris, H. S. Hoffman, and B. C. Griffith. (1957) The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* 54, 358-368.
- Liberman, A. M., I. G. Mattingly, and M. T. Turvey. (1972) Language codes and memory codes. In Coding Processes in Human Memory, ed. by A. W. Melton and E. Martin. (Washington, D. C.: V. H. Winston) 307-334.
- Lieberman, P. (1968) Primate vocalizations and human linguistic ability. *J. Acoust. Soc. Amer.* 44, 1574-1584.
- Lieberman, P. (1969) On the acoustic analysis of primate vocalizations. *Behav. Res. Meth. Instrum.* 1, 169-174.
- Lieberman, P. and E. S. Crelin. (1971) On the speech of Neanderthal man. *Ling. Inq.* 2, 203-222.
- Lieberman, P., E. S. Crelin, and D. H. Klatt. (1972) Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man and chimpanzee. *American Anthropologist* 74, 287-307.
- Lieberman, P., D. H. Klatt, and W. A. Wilson. (1969) Vocal tract limitations of the vocal repertoires of Rhesus monkey and other nonhuman primates. *Science* 164, 1185-1187.
- Lisker, L., F. S. Cooper, and A. M. Liberman. (1962) The uses of experiment in language description. *Word* 18, 83-106.
- Mattingly, I. G. (1968) Synthesis by rule of General American English. Supplement to Haskins Laboratories Status Report on Speech Research, 1-223.

- Mattingly, I. G. (1971) Synthesis by rule as a tool for phonological research. *Lang. Speech* 14, 47-56.
- Mattingly, I. G. (1972) Speech cues and sign stimuli. *American Scientist* 60, 327-337.
- Mattingly, I. G. and A. M. Liberman. (1969) The speech code and the physiology of language. In Information Processing in the Nervous System, ed. by K. N. Leibovic. (New York: Springer Verlag) 97-117.
- Mattingly, I. G., A. M. Liberman, A. K. Syrdal, and T. Halwes. (1971) Discrimination in speech and nonspeech modes. *Cog. Psychol.* 2, 131-157.
- Milner, B. (1954) Intellectual functions of the temporal lobe. *Psychol. Bull.* 51, 42-62.
- Milner, B. (1958) Psychological defects produced by temporal lobe excision. *Proceedings of the Association for Research of Nervous and Mental Disorders* 36, 244-257.
- Milner, B. (1967) Brain mechanisms suggested by studies of temporal lobes. In Brain Mechanisms Underlying Speech and Language, ed. by F. L. Darley. (New York: Grune and Stratton) 122-132.
- Nye, P. (1968) Research on reading aids for the blind--a dilemma. *Med. Biolog. Eng.* 6, 43-51.
- Orr, D. B., H. L. Friedman, and J. C. C. Williams. (1965) Trainability of listening comprehension of speeded discourse. *J. Educ. Psychol.* 56, 148-156.
- Pisoni, D. (1971) On the nature of categorical perception of speech sounds. Doctoral dissertation, University of Michigan. (Reproduced as Supplement to Haskins Laboratories Status Report on Speech Research.)
- Semmes, J. (1968) Hemispheric specialization: A possible clue to mechanism. *Neuropsychologia* 6, 11-27.
- Shankweiler, D. (1971) An analysis of laterality effects in speech perception. In The Perception of Language, ed. by D. L. Horton and J. J. Jenkins. (Columbus, Ohio: Chas. E. Merrill) 185-200.
- Shankweiler, D. and M. Studdert-Kennedy. (1967) Identification of consonants and vowels presented to left and right ears. *Quart. J. Exp. Psychol.* 19, 59-63.
- Sperry, R. W. and M. S. Gazzaniga. (1967) Language following surgical disconnection of the hemispheres. In Brain Mechanisms Underlying Speech and Language, ed. by C. H. Millikan and F. L. Darley. (New York: Grune and Stratton) 108-121.
- Stevens, K. N. (1960) Toward a model for speech recognition. *J. Acoust. Soc. Amer.* 32, 47-55.
- Stevens, K. N. and M. Halle. (1967) Remarks on analysis by synthesis and distinctive features. In Models for the Perception of Speech and Visual Form, ed. by W. Wathen-Dunn. (Cambridge, Mass.: MIT Press) 88-102.
- Stevens, K. N., A. M. Liberman, S. E. G. Ohman, and M. Studdert-Kennedy. (1969) Cross-language study of vowel perception. *Lang. Speech* 12, 1-23.
- Studdert-Kennedy, M. and F. S. Cooper. (1966) High-performance reading machines for the blind; psychological problems, technological problems, and status. Proceedings of Saint Dunstan's Conference on Sensory Devices for the Blind. (London) 317-342.
- Studdert-Kennedy, M. and D. Shankweiler. (1970) Hemispheric specialization for speech perception. *J. Acoust. Soc. Amer.* 48, 579-594.
- Studdert-Kennedy, M., D. Shankweiler, and D. Pisoni. (1972) Auditory and phonetic processes in speech perception: Evidence from a dichotic study. *Cog. Psychol.* 3, 455-466.

- Teuber, H.-L. (1960) Perception. In Handbook of Physiology Section 1: Neurophysiology, Volume III, ed. by J. Field. (Washington, D. C.: American Physiological Society) 1595-1669.
- Umilta, C., N. Frost, and R. Hyman. (1972) Interhemispheric effects on choice reaction times to one-, two-, and three-letter displays. *J. Exp. Psychol.* 93, 198-204.
- Vinegrad, M. (1970) A direct magnitude scaling method of investigating categorical versus continuous modes of speech perception. Haskins Laboratories Status Report on Speech Research SR-21/22, 147-156.
- Warren, R. M., C. J. Obusek, R. M. Farmer, and R. T. Warren. (1969) Auditory sequence: Confusion of patterns other than speech or music. *Science* 164, 586-587.
- Wood, C. C. (in preparation) Levels of processing in speech perception: Neurophysiological and cognitive analyses. Unpublished Ph.D. dissertation, Yale University.
- Wood, C. C., W. R. Goff, and R. S. Day. (1971) Auditory evoked potentials during speech perception. *Science* 173, 1248-1251.