ON THE RELATIONSHIP OF SPEECH TO LANGUAGE

JAMES E. CUTTING

Wesleyan University

JAMES F. KAVANAGH

National Institute of Child Health and Human Development

The relationship of speech to language has yet to be fully understood. In hope of elucidating this relationship we present a framework which considers speech and language as separate entities in a symbiotic partnership, performing similar functions toward similar ends but at different levels. Intuitive and logical support is given for their separation, and we consider the possible gains in the understanding of language by taking into account that it is usually spoken. Three approaches to the understanding of the role of speech in language are then pursued. The first is to compare the rules of speech (phonology) with those of language (syntax and semantics taken together); the second is to compare the development of speech in man and in the child, and the third is to compare sign language with speech. Given the power of these approaches, we must also remember the holism of the speech-language system.

WHAT IS THE RELATIONSHIP OF SPEECH TO LANGUAGE?

To some, this question may seem odd. One reason may be that, just as with the terms null and void, laypersons and scientists alike often view the terms speech and language as duplications of one another. At second glance one realizes that this is not true: speech could be considered as the spoken vehicle of language. This view would seem to place speech inside language, giving it the same relationship as the part to the whole.

Only recently have speech scientists, psychologists, linguists, anthropologists, and philosophers, among others, begun to look in earnest beyond these first and second glances; only recently have they begun to treat speech and language as separate entities in a symbiotic partnership. This third view, just as the previous ones, may not be entirely correct, but it has considerable intuitive and empirical support. Moreover, it provokes some interesting questions. For example, if language and speech are independent it must be possible to have language without speech and speech without language.

LANGUAGE WITHOUT SPEECH

There are several contenders for the label language without speech. Many are controversial. Consider first the sign languages of the deaf, particularly American Sign Language (ASL). This mode of communication uses hand gestures in relationship to the head and torso, along with much eye contact, to convey meaning from signer to sign receiver. Clearly there is no speech in ASL, no tongue movements to shape sound. This, among other features of sign languages, has led some researchers to question whether ASL is, indeed, a language at all. The title of Hans Furtth's book Thinking Without Language (1966) bespeaks this position; Bellugi and Klima's forthcoming book The Signs of Language (in press), on the other hand, will have a different view. Rather than enter into this debate, which may be more acrimonious than fruitful, some have chosen to observe how sign languages differ from spoken languages. We shall return to these observations in some detail.

Another illustration of language without speech is seen in certain cases of congenital anarthria, where the patient never acquires the ability to speak but can understand language easily. Christy Brown, for example, grew up with little speech, but had language abilities refined enough to write a best-seller, Down All the Days (1970). In an even more extreme example, Lenneberg (1962) reports the case of a child who had no speech, but could understand language nearly as well as his unafflicted age-mates.

A third possibility of language without speech is the most controversial and concerns the considerable efforts undertaken to teach language to chimpanzees. It is clear that chimps cannot learn to talk even given the most extensive training: their vocal tract simply appears to be inadequate (Lieberman, Crelin, and Klatt, 1972). They can, however, become remarkably adept at using the sign gestures of ASL (Gardner and

*JAMES E. CUTTING, Ph.D., is an assistant professor in the Department of Psychology, Wesleyan University, Middletown, Connecticut, and a research associate with Haskins Laboratories, New Haven, Connecticut. JAMES F. KAVANAGH, Ph.D., is acting program director, Growth and Development Branch, National Institute of Child Health and Human Development, Bethesda, Maryland.

Reprinted from ASHA
August 1975, Vol. 17, No. 8
Gardner, 1969; Fouts, 1973), manipulating plastic symbols on a magnetized board to convey meaning (Premack, 1971), or "reading and sentence completion" of computer-displayed geometric symbols (Rumbaugh, Gill, and von Glaserfeld, 1973). Are chimps capable of language behavior, or merely language-like behavior? Fodor, Bever, and Garrett (1974) remain unconvinced that these demonstrations are even relevant to language; Lieberman (1973), on the other hand, finds them compelling. This is another controversy we choose to avoid. Regardless of whether chimps do or do not have language, we think it useful to observe what chimpanzees can and cannot do for the purpose of investigating the scope of language without speech.

**SPEECH WITHOUT LANGUAGE**

There are also several contenders for the label *speech without language*. Again, some are controversial. The early babbling of the infant is often thought to be nonlinguistic (Jakobson, 1968; Kewley-Port and Preston, 1974). Brain-damaged patients with extreme forms of expressive aphasia often speak with good rhythm and intonation patterns, but with no apparent words or meaning (Green, 1973). The "speaking in tongues," or glossolalia, often associated with Pentecostal churches has been found to lack underlying structures necessary in more worldly languages (Samarin, 1972). Some consider all three of these examples more akin to song than to language, and, indeed, *glossa* means tongue and *lalita* lullaby. One can avoid any controversy, however, by looking to song lyrics themselves for examples of speech without language. The "fa-la la-la-la" of certain Christmas carols and the "sha-boom sha-boom" of certain popular songs of the 1950s and 1960s are surely acceptable to critics as lacking linguistic content. These are speech sounds for sound's sake. They have no duality of patterning so familiar to spoken languages (Hockett and Altmann, 1958); that is, they are sound without meaning.¹

¹The conference was entitled "Communication by Language—The Role of Speech in Language." Those who attended or contributed to the conference included, in addition to the present authors, Ursula Bellugi, James F. Bosma, Peter D. Elmas, Jerry A. Fodor, Gordon W. Hewes, Ira J. Hirsh, Janelen Huttenlocher, James J. Jenkins, R. Paul Kiparsky, Edward S. Klima, Alvin M. Lieberman (cochairman with Kavanagh), Philip Lieberman, Peter Marler, Ignatius C. Mattingly, David S. Palermo, David Premack, Peter C. Reynolds, John Robert Ross, Robert E. Shaw, William C. Stokoe, Jr., and Michael Studdert-Kennedy. The conference proceedings are published by the MIT Press as *The Role of Speech in Language* (Kavanagh and Cutting, 1975).

²We have borrowed the notion of a speech chain from Denes and Pinson (1973)—which includes the vocal tract, air vibrations, and the ear—and extended it to include intellect at both ends. The result could still be called the speech chain, but we propose to substitute the vocal tract and ear with the hands and eye, respectively, when dealing with sign language, and to substitute human intellect with that of chimpanzees and even birds when dealing with animal communication. The end result can only be considered the communication chain.

**A FRAMEWORK FOR THE STUDY OF SPEECH AND LANGUAGE**

If speech and language are as isolable from one another as they appear to be in the above examples, many interesting questions arise. How do speech and language function in concert, and, more particularly, what are the effects of one upon the other? In October 1973, a group of researchers, many of whom are directly involved in the controversies mentioned earlier, met under sponsorship of the National Institute of Child Health and Human Development at Columbia, Maryland, for three days of presentations and discussions. Their topic was the role of speech in language, and the following is, in part, based on those proceed-
works upward in the communication chain to constrain and alter language and perhaps even intellect; language, working in the reverse direction, exerts downward constraints to alter speech, the vocal tract, and perhaps the ear as well. Evidence for evolutionary change in the shape of the mind is difficult to come by. Evidence for evolutionary change in the shape of the vocal tract, however, can be seen by comparing fossil skulls of certain hominids with those of modern man. Philip Lieberman, at the conference and in previous publications, suggested that the human vocal tract assumed its present configuration specifically to make speech possible. This view is contrary to the more venerated notion that speech is merely a faculty overlaid on eating and respiratory functions. Evidence that the newer, evolutionary view is correct stems partly from the fact that man, in addition to being the only creature to speak, may be the only creature to choke easily on his food. While these downward constraints on the vocal tract are important, it is the upward constraints, those which shape language and the mind, which are perhaps the more interesting changes in evolution, and it is those which are more directly relevant to the role of speech in language.

Three approaches seem relevant to our goal of understanding the relationship of speech to language. First, one can focus on speech itself, or more specifically on phonology, to obtain insights about the workings of language and of the mind. Second, one can trace inasmuch as possible the development of speech in man and child, making inferences about language and intellect behind the expansion of ability in vocal communication. Third, one can look at the linguistic structures of sign language, the most important form of language without speech, with an eye toward differences between sign and speech and how they affect the more abstract levels of the communication chain.

PHONOLOGY AND THE LANGUAGE OF THE MIND

Speech scientists and linguists have always treated speech and language as separate entities. Their problem, according to R. Paul Kiparsky and John Robert Ross, who attended the conference, is a failure to map out, in a nontrivial manner, the functional and structural relationship between them. One way to accomplish this appears to be to observe interactions of phonology and semantax. For example, “John’s in Boston” is a perfectly good sentence. “Bill’s happier in Portland than John’s in Boston,” however, is not. In this example by Kiparsky the phonology of the phrase “John is in Boston” is dictated by higher level rules—mind shapes mouth. Are there examples of mouth shaping mind, where phonological rules dictate semantic structure? Perhaps, but they appear much more difficult to find at present.

A second way to accomplish our goal, then, is to draw parallels between phonological and semantactic grammars. Ross outlined several parallels, one of which might be termed a simplification process at both levels. At the semantactic level speakers tend to reduce complex sentences to simple ones. Rather than saying “I know someone who is tall,” for example, one is more likely to say a shorter and simpler sentence “I know someone tall.” At the phonological level speakers tend to reduce multisyllable utterances into one or two syllable utterances, especially when among friends. Thus, “Did you eat yet?” is easily shortened to “Did y’eat yet?” and finally to “I eat yet.” There are, however, problems with such parallels. Just as the correlation does not imply causation in statistical analysis, parallels between phonology and semantics do not necessarily imply upward or downward constraints in the communication chain. Nevertheless, such groundwork is vital to the field if it is to become ripe for new discoveries.

DEVELOPMENT OF SPEECH IN MAN AND CHILD

We can sketch only some of the more important and interesting issues in this awesomely broad, second approach. One issue, for example, is why speech developed so late in man—perhaps only 50,000 years ago—and develops so late in the child—between one and two years. One reason for this “lateness is directly related to functional anatomy, as suggested earlier. Lieberman reconstructed from fossil remains the vocal tracts of premodern man and compared them to those of modern adults and neonates. Of the three, the vocal tracts of premodern man and the modern neonate were most similar and lacked the particular shape requisite for full-range speech sounds of the modern adult. Thus, ontogeny recapitulates phylogeny, and one answer to the lateness problem in both man and child appears to be physiological inadequacy. Physiology, however, cannot be the entire answer. The child’s vocal tract becomes adequate many months before speech is produced in a regular fashion. By inference, this may have been true for premodern man as well. Therefore, other factors such as cognitive ability must be considered: men and children need something to say as well as the apparatus with which to say it.

The tardiness yet pervasiveness of speech seems paradoxical. Whereas language without speech is thought by some to be impoverished, language abilities may develop before speech abilities. Gordon Hewes (1973), for example, has suggested that language first developed in prehistory through the use of gestures perhaps similar to those of modern sign languages. William Stokoe, at the conference, claimed that sign language develops in the deaf child before speech develops in the normal child. These notions, if true, would seem to indicate that sign is more “natural” to language than is speech, an irony indeed. The resolution of this apparent paradox may be to assume that speech and language evolved separately, perhaps at separate times, and only later coevolved into a more or less unified and symbiotic system. The independent
evolution of speech is supported by Mattingly (1972). He noted structural parallels between speech, certainly the most complex signaling system in nature, and various rudimentary animal communication systems which could hardly be called language or even languagelike.

If language by sign developed earlier than speech, or at least independent of it, why did speech supplant sign as the major vehicle of language? Surely the answer must be more complex than to free the hands for manual skills such as hunting, gathering, toolmaking, and cooking. One reason, we can safely assume, concerns speed of communication. Ursula Bellugi noted at the conference that modern sign languages are not as rapid as speech (see also Bellugi and Fischer, 1972). Proto-sign was surely no faster and could not compete with the more rapid, newly evolved form of communication. This view seems reasonable. Even speech is woefully slow at times. Rips of the tongue often reveal telescopic jumps where speech skips ahead many syllables as if to catch up with the more nimble leaps of the mind. There may be evolutionary and ever-present pressures to speed up communication. Perhaps sign lost out to speech because of them.

Another reason for the change from sign to speech may be related to modality. Put in its simplest form, almost all objects in nature are opaque to the eye, few are "opaque" to the ear; that is, one cannot see through foliage and rocks, but he or she can hear "through" or at least around them. This feature becomes vitally important when one walks or runs through dense jungles and high grasses, as did man's forebears, where vision is often very restricted. In this light, it is necessary to consider the role of vocalizations in animal communication, comparing them to the role of speech in language. Two types of creatures are of particular interest: primates, because of their evolutionary relationship to man, and songbirds, because of impressive analogs between the acquisition of birdsong and of speech.

Peter Marler told the conference about comparative ethological trends in Asian and African primates which are relevant to development of speech and language in man. As primates develop a more complex vocal repertory, they also tend to become more terrestrial (living on the ground rather than in trees), less territorial, and more inclined to live in large troops. All of these are trends toward the social state of man. More importantly, a major change of emphasis in communication appears to be correlated with this trend. With these other developments, the largest portion of signaling repertoires shifts from between-troop warning calls and vocal displays to within-troop social calls. Parallel to this change in type of communication is a change in "vocabulary," from a discrete and limited set of calls to a graded and less-bounded call system. This trend allows for a larger and more subtle repertory of vocal sounds. Marler interprets this move toward graded systems as approximations of speechlike behavior in man.

From a view external to that of the speech perceiver, Marler is correct: human speech is extremely graded. For example, if many samples of human speech were displayed on sound spectrograms and compared to each other, one would see an impressive dearth of discrete differences among the speech sounds. They would look, as Hockett (1955) has suggested, like so many smashed Easter eggs. To be sure, humans do not perceive speech in a graded or continuous manner; it seems to segment itself into syllables and phonemes almost automatically. How we accomplish the feat of reassembling the smashed eggs, the units of speech, remains largely a mystery, as those involved in the problem of machine recognition of speech can attest. Viewed from the "outside," then, as any computer or intelligent nonhuman must view speech, it is strikingly graded and continuous. This raises an interesting issue. Just as computers have difficulty segmenting human speech, humans have difficulty segmenting the graded calls of chimpanzees, which are necessarily viewed from the "outside." Do chimps and other primates segment their graded vocalizations? This is an important question. Whether they do or do not, however, the emphasis on the evolutionary role of speech in language might well be placed on perception rather than on production.

The prominence of perception over production receives support from birdsong, as well as from speech itself. Consider first the songs of passerine birds. The white-crowned sparrow, for example, must hear versions of his species-specific song if he is to produce it, and he must hear it during his first year, well before he begins to sing it. Furthermore, he must continue to hear himself and fellow white-crowned sparrows as he produces approximations to full song during the following year. Surgical deafening at any time before the advent of full song inhibits the production process and full song will not develop. In an analogous fashion humans may need to perceive speech before they can start to produce it, and later they may need to compare their productions with those of adults before speech becomes regularized. Critical periods for humans are probably much less inflexible than for songbirds, but a parallel is unmistakable. Evidence suggests that infants can perceive speech-relevant sounds well before they can produce them. Peter Elmas, at the conference and in previous work (Elmas et al., 1971), presented data that one-month-old infants are able to discriminate phonetically relevant features in computer-generated tokens of speech much better than similar but phonetically irrelevant features. These discriminations, which are requisites for speech segmentation, occur at least a year before the same phonetic distinctions will be accurately produced (Kewley-Port and Preston, 1974).

If one considers speech as a "species-specific song" in a broad sense, infants must be exposed to elements of the "full song" long before they can produce it. Infants deaf from birth have extreme difficulty in acquiring speech, but children who become deaf later,
at age five or 10, for example, may continue to have remarkably normal speech for the rest of their lives, just as the white-crowned sparrow deafened after the development of song in his second year will continue to sing in a normal manner.

In addition, like humans, white-crowned sparrows have dialects according to geographical region. These aspects of full song appear to be first learned through exposure long before the young bird ever sings. Recent research with humans has shown that young infants begin to learn by the age of two months the more exotic, "dialectic" aspects of their to-be-native language which two-month-old infants in other lands will not have learned (Streeter, 1974). Again, this is long before the sounds will be produced and used to convey meaning in spoken language.

Ontogenetic and phylogenetic observations about the acquisition of speech have gone well beyond our first approach to the role of speech in language, that of observing phonology itself. Yet, like that approach, this second one is still very new and has only recently begun to bear fruit. Evidence from the calling systems of primates and of songbirds, as well as that presented by Mattingly (1972), supports the view that speech has strong evolutionary ties independent of language. Thus far, however, we have presented little information about how speech as a signaling system was applied to language and what effects that application had. This is crucial to our goal of discerning the role of speech in language. Our third approach is addressed to this question, but necessarily in an indirect fashion.

**COMPARISONS OF SIGN LANGUAGE AND SPEECH**

If perception is a requisite for production of speech, as we have suggested earlier, what is the effect on language and intellect when that channel of perception is totally blocked? Robbed of audition from birth, the deaf human may have no opportunity to develop speech and may have to use the slower sign gestures to communicate. Some researchers have suggested that the choice of sign over speech may have intellectual costs. In some cases, however, it is clear that there are no such costs to deaf signers even when they are compared to normal speakers. But the question about the size or intellectual capacity of the mind should be separated from the question about the shape of the mind. The shape of a soundless language and the intellect behind it is the issue addressed by Bellugi, Klima, and Stokoe, among others.

Aside from the sheer scope of trying to compare all of sign to all of speech, there are several other problems. One is data base. Only one person in 1000 is deaf, and only one deaf person in 10 is the child of deaf parents. Thus, it is only one child in 10,000 who learns sign as a native language. The other nine in 10,000 will probably learn sign, but in conjunction with speech which might "contaminate" the study of pure sign. Secondly, there is the problem of the pervasive influences of the spoken culture around enclaves of native signers. In America, among signers of ASL, there are at least three forms of signs: (1) finger-spelled words of English which may not have a direct analog in sign, (2) signed English which is an approximation of English morphology and syntax, and (3) natural sign. Native signers typically use all three, but it is only the latter which is of primary interest here. Thirdly, there are differences between sign and pantomime which must be closely observed. Sign is only partially iconic whereas pantomime is almost exclusively so. The icon, or visual image, is often drawn or shown with the fingers and hands in front of the signer/pantomimist and referred to later in the sign/pantomime discourse. With all these complexities it becomes evident that any effort to study sign language by the nonsigning researcher is difficult without the aid of native-signing collaborators. Stokoe at Gallaudet College in Washington, D. C., and Bellugi and Klima at the Salk Institute in California rely heavily on their deaf colleagues.

Comparing sign to speech, one first finds that sign has no sounds, no phones, and no "phonology" in the normal sense. Are there meaningless units in sign? Yes, but they do not correspond exactly to the phoneme or even to the syllable. The three important features of a sign, in a psychological sense, appear to be the hand configuration, the place of articulation of the designating hand with respect to the head, torso, or other hand, and the movement of the hand once it is there. Each configuration, place, and movement is meaningless in itself just as phones are meaningless. It may seem ironic that meaninglessness is important to communication; one could easily have predicted the opposite. Nevertheless, it is the combination of such units which makes meaningful words and signs possible. Some combinations are easier than others to produce, and some, while easy to articulate, simply seem wrong: whereas *bnick* (to use an example from Klima) is easy to pronounce, it does not conform to English phonology. Thus, phonological rules constrain the possible combinations of phones. There are sign analogs to *bnick*. Certain hand configurations seem wrong to native signers when accompanied by certain movements or coupled with a certain place of articulation. In the broadest sense, then, sign has a "phonology" analogous to that of any spoken language. When comparing the influences of sign and of speech on language and intellect one must remember that he is not comparing systems in the presence or absence of phonology, but rather systems with different phonologies. This makes our task all the more difficult, but all the more intriguing as well.

From this brief look at the phonology of sign it may appear to consist merely of articulatory dos and don'ts. This is incorrect. The phonology of sign, if we may use the phrase, is broader than that. Perhaps more important than articulation rules are the temporal constraints alluded to earlier. Since speech is faster than
sign, sign must somehow try to catch up. Bellugi and Fischer (1972) asked how sign saves time and still communicates unambiguously. The answers fall into at least three categories: doing without, incorporation, and bodily or facial shifts. Doing without often means simply doing without the redundancy of spoken and written language. Bellugi and Fischer note that the signed version of the complex sentence “John likes Mary, so he goes and visits her a lot, and he often takes her out to dinner, though sometimes he cooks for her” would scan (when translated back into English) something like “John like Mary, well, go visit much, often take out eat, but sometimes cook for.” Clearly, much has been dropped in the signed version, but the message is essentially identical. Incorporation, the second way to short-cut in sign, takes many forms. Often sign incorporates iconic spatial references. A simple example would be to compare the two signed sentences corresponding to “She is bigger than me” and “She is much bigger than me.” Both signed sentences would take the same amount of time to pronounce but in the second form the sign for large (bigger than) would be exaggerated. Bodily and facial shifts, the third major class of sign accelerators, deliver information in parallel with the sign discourse. For example, the hand gestures corresponding to the sentences “I know that” and “I don’t know that” are identical. The signed version of the second sentence is accompanied by a headshake, or a small frown, indicating negation, thus saving time. Bellugi and Fischer do not claim that this small list includes all time-saving devices in sign, but it is interesting that these three—doing without, incorporation, and bodily shifts—are exactly those which make face-to-face verbal communication so much easier and faster than communication by telephone. Furthermore, they are exactly the reasons that conferences and meetings, where people are often drawn together from great distance and at great expense, are more prevalent and more rewarding than conference telephone calls, even though the latter may be cheaper.

Systematic comparisons of sign and speech have only just begun. Much of the present research may look like so much dabbling, but underlying it is the need for asking the right questions, which cannot be posed until we have dabbled. Promising avenues have been started by Bellugi, Klima, and others, and a few deserve mention here. First, just as there are slips of the tongue in speech, there are “slips of the hand” in sign. Fromkin (1973) has analyzed these faux pas in speech and found richly rewarding insights into the serial organization of speech. Studies of slips of the hand will be equally rewarding in unraveling the structures of sign. Second, just as there are infantile or “baby talk” forms of speech, there are infantile forms of sign. In some ways these are similar to speech, in others they are different. The acquisition of signs by children is certainly worthy of study to the extent that, for instance, Brown (1973) has studied the first spoken sentences of normal children. Third, psychologists have been interested in the different types of forgetting that occur for information presented by eye and information by ear. Typically these memory errors are different, particularly with regard to most recently occurring items in a list. Bellugi has found evidence that sign receivers forget lists of words in a manner nearly identical to the way normal listeners forget lists of words which are spoken, but not in the manner normals forget those words when written. By extension, perceiving sign may be more similar to listening to speech than to reading, even though both sign receiving and reading are visual skills.

**HOLISM OF SPEECH AND LANGUAGE**

A word of caution must be inserted at this point. While it is clear that speech and language can be logically separated, whether by comparing phonology and semantics, by postulating their separate genetic developments, or by comparing language with and without speech, they remain part of one system. James Jenkins and Robert Shaw, playing devil’s advocates at the conference, saw a danger in the fractionation of speech and language and subsequent overanalyses which may follow. As a historical case in point, they noted how the field of aphasia research has suffered from this division. After reviewing 50 years of empirical research on large samples of brain-damaged patients, they found few, if any, examples of pure productive aphasia (language without speech) or pure receptive aphasia (speech without language).

Perhaps the third view of the relationship between speech and language, that they are separate entities in a symbiotic partnership, should be tempered. Separateness may imply an independence which surely does not exist in the normal speech-language-communication system in man. Accepting this cautionary note, exploration into the relationship of speech to language has only just begun and should prove a fascinating and fruitful line of research for those in many scientific disciplines which converge on communication in man.

**REFERENCES**