The Origin of Speech

by Charles F. Hockett

September 1960

Man is the only animal that can communicate by means of abstract symbols. Yet this ability shares many features with communication in other animals, and has arisen from these more primitive systems.

About 50 years ago the Linguistic Society of Paris established a standing rule barring from its sessions papers on the origin of language. This action was a symptom of the times. Speculation about the origin of language had been common throughout the 19th century, but had reached no conclusive results. The whole enterprise in consequence had come to be frowned upon—as futile or crackpot—in respectable linguistic and philological circles. Yet amidst the speculations there were two well-reasoned empirical plans that deserve mention even though their results were negative.

A century ago there were still many corners of the world that had not been visited by European travelers. It was reasonable for the European scholar to suspect that beyond the furthest frontiers there might lurk half-men or man-apes who would be "living fossils" attesting to earlier stages of human evolution. The speech (or quasi-speech) of these men (or quasi-men) might then similarly attest to earlier stages in the evolution of language. The search was vain. Nowhere in the world has there been discovered a language that can validly and meaningfully be called "primitive." Edward Sapir wrote in 1921: "There is no more striking general fact about language than its universality. One may argue as to whether a particular tribe engages in activities that are worthy of the name of religion or of art, but we know of no people that is not possessed of a fully developed language. The lowliest South African Bushman speaks in the forms of a rich symbolic system that is in essence perfectly comparable to the speech of the cultivated Frenchman."

The other empirical hope in the 19th century rested on the comparative method of historical linguistics, the discovery of which was one of the triumphs of the period. Between two languages the resemblances are sometimes so extensive and orderly that they cannot be attributed to chance or to parallel development. The alternative explanation is that the two are divergent descendants of a single earlier language. English, Dutch, German and the Scandinavian languages are related in just this way. The comparative method makes it possible to examine such a group of related languages and to construct, often in surprising detail, a portrayal of the common ancestor, in this case the proto-Germanic language. Direct documentary evidence of proto-Germanic does not exist, yet understanding of its workings exceeds that of many languages spoken today.

There was at first some hope that the comparative method might help determine the origin of language. This hope was rational in a day when it was thought that language might be only a few thousands or tens of thousands of years old, and when it was repeatedly being demonstrated that languages that had been thought to be unrelated were in fact related. By applying the comparative method to all the languages of the world, some earliest reconstructible horizon would be reached. This might not date back so early as the origin of language, but it might bear certain earmarks of primitiveness, and thus it would enable investigators to extrapolate toward the origin. This hope also proved vain. The earliest reconstructable stage for any language family shows all the complexities and flexibilities of the languages of today.

These points had become clear a half-century ago, by the time of the Paris ruling. Scholars cannot really approve of such a prohibition. But in this instance it had the useful result of channeling the energies of investigators toward the gathering of more and better information about languages as they are today. The subsequent progress in understanding the workings of language has been truly remarkable. Various related fields have also made vast strides in the last half-century: zoologists know more about the evolutionary process, anthropologists know more about the nature of culture, and so on. In the light of these developments there need be no apology for reopening the issue of the origins of human speech.

Although the comparative method of linguistics, as has been shown, throws no light on the origin of language, the investigation may be furthered by a comparative method modeled on that of the zoologist. The frame of reference must be such that all languages look alike when viewed through it, but such that within it human language as a whole can be compared with the communicative systems of other animals, especially the other hominoids, man's closest living relatives, the gibbons and great apes. The useful items for this sort of comparison cannot be things such as the word for "sky"; languages have such words, but gibbon calls do not involve words at all. Nor can they be even the signal for "danger," which gibbons do have. Rather, they must be the basic features of design that can be present or absent in any communicative system, whether it be a communicative system of humans, of animals or of machines.

With this sort of comparative method it may be possible to reconstruct the communicative habits of the remote ancestors of the hominoid line, which may be called the protohominoids. The task, then, is to work out the sequence by
which that ancestral system became lan-
guage as the hominids—the man-apes
and ancient men—became man.

A set of 13 design-features is pre-
sented in the illustration on the op-
posite page. There is solid empirical jus-
tification for the belief that all the lan-
guages of the world share every one of
tem. At first sight some appear so trivial
that no one looking just at language
would bother to note them. They become
worthy of mention only when it is real-
ized that certain animal systems—and
certain human systems other than lan-
guage lack them.

The first design-feature—the "vocal-
auditory channel"—is perhaps the most
obvious. There are systems of communica-
tion that use other channels, for ex-
ample, gesture, the dancing of bees or the
courtship ritual of the stickleback. The
vocal-auditory channel has the advan-
tage—at least for primates—that it leaves
much of the body free for other activities
that can be carried on at the same time.

The next two design-features—"rapid
fading" and "broadcast transmission and
directional reception," stemminng from
the physics of sound—are almost un-
avoidable consequences of the first. A
linguistic signal can be heard by any
auditory system within earshot, and the
source can normally be localized by bin-
aural direction-finding. The rapid fad-
ing of such a signal means that it does
not linger for reception at the hearer's
convenience. Animal tracks and spoors,
on the other hand, persist for a while; so
of course do written records, a product
of man's extremely recent cultural evo-

The significance of "interchangeabil-
ity" and "total feedback" for language
becomes clear upon comparison with
other systems. In general a speaker of a
language can reproduce any linguistic
message he can understand, whereas the
characteristic courtship motions of the
male and female stickleback are differ-
ent, and neither can act out those ap-
propriate to the other. For that matter
in the communication of a human moth-
er and infant neither is apt to transmit
the characteristic signals in manifest
the typical responses of the other. Again,
the speaker of a language hears, by total
feedback, everything of linguistic rele-
ance in what he himself says. In con-
trast, the male stickleback does not see
the colors of his own eye and belly that
are crucial in stimulating the fe-
males. Feedback is important, since it
makes possible the so-called internali-
zation of communicative behavior that
constitutes at least a major portion of
"thinking."

The sixth design-feature, "specializa-
tion," refers to the fact that the bodily
effort and spreading sound waves of
speech serve no function except as sig-
nals. A dog, panting with his tongue
hanging out, is performing a biologically
essential activity, since this is how dogs
cool themselves off and maintain the
proper body temperature. The panting
dog incidentally produces sound, and
thereby may inform other dogs (or hu-
mans) as to where he is and how he
feels. But this transmission of informa-
tion is strictly a side effect. Nor does the
dog's panting exhibit the design-feature of"semantics." It is not a signal mean-
ing that the dog is hot; it is part of being
hot. In language, however, a message
triggers the particular result it does be-
cause there are relatively fixed associa-
tions between elements in messages (e.g.,
words) and recurrent features or situa-
tions of the world around us. For
example, the English word "salt" means
salt, not sugar or pepper. The calls of
gibbons also possess sematicity. The
gibbon has a danger call, for example,
and it does not in principle matter that
the meaning of the call is a great deal
broader and more vague than, say, the
cry of "Fire!"

In a semantic communicative system
the ties between meaningful message-
elements and their meanings can be ar-
bitrary or nonarbitrary. In language the
ties are nonarbitrary. The word "salt" is
not salty nor granular; "dog" is not "canine;"
"whale" is a small word for a large ob-
ject; "microorganism" is the reverse. A
picture, on the other hand, looks like
what it is a picture of. A bee dances
faster if the source of nectar she is re-
porting is closer, and slower if it is far-
ther away. The design-feature of arbi-
trariness has the disadvantage of being
arbitrary, but the great advantage that
there is no limit to what can be com-
municated about.

Human vocal organs can produce a
huge variety of sound. But in any one
language only a relatively small set of
ranges of sound is used, and the differ-
ences between these ranges are function-
ally absolute. The English words "pin"
and "bin" are different to the ear only at
one point. If a speaker produces a syl-
able that deviates from the normal pro-
mununciation of "pin" in the direction of
that of "bin," he is not producing still a
third word, but just saying "pin" (or
perhaps "bin") in a noisy way. The
hearer compensates if he can, on the
basis of context, or else fails to un-
derstand. This feature of "discreteness" in
the elementary signaling units of a lan-
guage contrasts with the use of sound
effects by way of vocal gesture. There is
an effectively continuous scale of de-
grees to which one may raise his voice
as in anger, or lower it to signal confi-
dentiality. Bee-dancing also is continu-
ous rather than discrete.

Man is apparently almost unique in
being able to talk about things that are
remote in space or time (or both) from
where the talking goes on. This feature—
displacement—seems to be definitely
lacking in the vocal signaling of man's
closest relatives, though it does occur in
bee-dancing.

One of the most important design-
features of language is "productIVITY":
that is, the capacity to say things that
have never been said or heard before
and yet to be understood by other speak-
ers of the language. If a gibbon makes
any vocal sound at all, it is one or an-
other of a small finite repertory of fa-
miliar calls. The gibbon call system can
be characterized as closed. Language is
open, or "productive," in the sense that
one can coin new utterances by putting
together pieces familiar from old utte-
rances, assembling them by patterns of
arrangement also familiar in old utter-
rances.

Human genes carry the capacity to
acquire a language, and probably also
a strong drive toward such acquisition,
but the detailed conventions of any one
language are transmitted extragenetical-
ly by learning and teaching. To what
extent such "traditional transmission" is
played a part in gibbon calls or for other
mammalian systems of vocal signals is
not known, though in some instances the
uniformity of the sounds made by a spe-
cies, wherever the species is found over
the world, is so great that genetics must
be responsible.

The meaningful elements in any lan-
guage—"words" in everyday parlance," morphemes" to the linguist—constitute
an enormous stock. Yet they are repre-
sented by small arrangements of a rel-
tively very small stock of distinguishable
sounds which are in themselves wholly
meaningless. This "duality of patterning" is illustrated by the English words

THIRTEEN DESIGN-FEATURES of ani-
mal communication, discussed in detail in
the text of this article, are symbolized on
opposite page. The patterns of the words
"pin," "bin," "team" and "meat" were
recorded at Bell Telephone Laboratories.
The origin of modern Germanic languages, as indicated by this "family tree," was proto-Germanic, spoken some 2,700 years ago. Comparison of present-day languages has provided detailed knowledge of proto-Germanic, although no direct documentary evidence for the language exists. It grew, in turn, from the proto-Indo-European of 5000 B.C. Historical studies cannot, however, trace origins of language back much further in time.

It should be noted that some of these 13 design-features are not independent. In particular, a system cannot be either arbitrary or nonarbitrary unless it is semantic, and it cannot have duality of patterning unless it is semantic. It should also be noted that the listing does not attempt to include all the features that might be discovered in the communicative behavior of this or that species, but only those that are clearly important for language.

It is probably safe to assume that nine of the 13 features were already present in the vocal-auditory communication of the protohominoids—just the nine that are securely attested for the gibbons and humans of today. That is, there were a dozen or so distinct calls, each the appropriate vocal response (or vocal part of the whole response) to a recurrent and biologically important type of situation: the discovery of food, the detection of a predator, sexual interest, need for maternal care, and so on. The problem of the origin of human speech, then, is that of trying to determine how such a system could have developed the four additional properties of displacement, productivity and full-blown traditional transmission. Of course the full story involves a great deal more than communicative behavior alone. The development must be visualized as occurring in the context of the evolution of the primate horde into the primitive society of food-gatherers and hunters, an integral part, but a part, of the total evolution of behavior.

It is possible to imagine a closed system developing some degree of productivity, even in the absence of the other three features. Human speech exhibits a phenomenon that could have this effect, the phenomenon of "blending." Sometimes a speaker will hesitate between two words or phrases, both reasonably appropriate for the situation in which he is speaking, and actually say something that is neither wholly one nor wholly the other, but a combination of parts of each. Hesitating between "Don't shout so loud" and "Don't yell so loud," he might come out with "Don't shell so loud." Blending is almost always involved in slips of the tongue, but it may
EVOLUTION OF LANGUAGE and some related characteristics are suggested by this classification of chordates. The lowest form of animal in each classification exhibits the features listed at the right of the class. Brackets indicate that each group possesses or has evolved beyond the characteristics exhibited by all the groups below. The 13 design-features of language appear in the colored rectangle. Some but by no means all of the characteristics associated with communication are presented in the column at right.
also be the regular mechanism by which a speaker of a language says something that he has not said before. Anything a speaker says must be either an exact repetition of an utterance he has heard before, or else some blended product of two or more such familiar utterances. Thus even such a smooth and normal sentence as "I tried to get there, but the car broke down" might be produced as a blend, say, of "I tried to get there but couldn't" and "While I was driving down Main Street the car broke down."

Children acquiring the language of their community pass through a stage that is closed in just the way gibbon calls are. A child may have a repertory of several dozen sentences, each of which, in adult terms, has an internal structure, and yet for the child each may be an indivisible whole. He may also learn new whole utterances from surrounding adults. The child takes the crucial step, however, when he first says something that he has not learned from others. The only way in which the child can possibly do this is by blending two of the whole utterances that he already knows.

In the case of the closed call-system of the gibbons or the protohominoids, there is no source for the addition of new unitary calls to the repertory except perhaps by occasional imitation of the calls and cries of other species. Even this would not render the system productive, but would merely enlarge it. But blending might occur. Let AB represent the food call and CD the danger call, each a fairly complex phonetic pattern. Suppose a protohominoid encountered food and caught sight of a predator at the same time. If the two stimuli were balanced just right, he might emit the calls ABCD or CDAB in quick sequence, or might even produce AD or CB. Any of these would be a blend. AD, for example, would mean "both food and danger." By

<table>
<thead>
<tr>
<th></th>
<th>A (SOME GRYLIDAE AND TETTIGONIIDAE)</th>
<th>B (BEE DANCING)</th>
<th>C (STICKLEBACK COURTSHIP)</th>
<th>D (WESTERN MEADOWLARK SONG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>THE VOCAL-AUDITORY CHANNEL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUDITORY, NOT VOCAL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>BROADCAST TRANSMISSION AND DIRECTIONAL RECEPTION</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>3</td>
<td>RAPID FADING (TRANSITORINESS)</td>
<td>YES, REPEATED</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>INTERCHANGEABILITY</td>
<td>LIMITED</td>
<td>LIMITED</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>TOTAL FEEDBACK</td>
<td>YES</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>SPECIALIZATION</td>
<td>YES</td>
<td>?</td>
<td>IN PART</td>
</tr>
<tr>
<td>7</td>
<td>SEMANTICITY</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>8</td>
<td>ARBITRARINESS</td>
<td>?</td>
<td>NO</td>
<td>IF SEMANTIC, YES</td>
</tr>
<tr>
<td>9</td>
<td>DISCREteness</td>
<td>?</td>
<td>YES</td>
<td>?</td>
</tr>
<tr>
<td>10</td>
<td>DISPLACEMENT</td>
<td>YES, ALWAYS</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>11</td>
<td>PRODUCTIVITY</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>12</td>
<td>TRADITIONAL TRANSMISSION</td>
<td>NO</td>
<td>PROBABLY NOT</td>
<td>?</td>
</tr>
<tr>
<td>13</td>
<td>DUALITY OF PATTERNING</td>
<td>(TIVIAL)</td>
<td>NO</td>
<td>?</td>
</tr>
</tbody>
</table>

EIGHT SYSTEMS OF COMMUNICATION process in varying degrees the 13 design features of language. Column A refers to members of the cricket family. Column B concerns only Western music since the time of Bach. A question mark means that it is
virtue of this, A, B, and C would acquire new meanings, respectively "food without danger" and "danger without food." And all three of these calls—A, B, C, and D—would now be composite rather than unitary, built out of smaller elements with their own individual meanings. A would mean "food," B, "no danger," C, "no food," and D, "danger."

But this is only part of the story. The generation of a blend can have no effect unless it is understood. Human beings are so good at understanding blends that it is hard to tell a blend from a rote repetition, except in the case of slips of the tongue and some of the earliest and most tentative blends used by children. Such powers of understanding cannot be ascribed to man's prehuman ancestors. It must be supposed, therefore, that occasional blends occurred over many tens of thousands of years (perhaps, indeed, they still may occur from time to time among gibbons or the great apes), with rarely any appropriate communicative impact on hearers, before the understanding of blends became speedy enough to reinforce their production. However, once that did happen, the earlier closed system had become open and productive.

It is also possible to see how faint traces of displacement might develop in a call system even in the absence of productivity, duality and thoroughgoing traditional transmission. Suppose an early hominid, a man-ape say, caught sight of a predator without himself being seen. Suppose that for whatever reason—perhaps through fear—he sneaked silently back toward others of his band and only a bit later gave forth the danger call. This might give the whole band a better chance to escape the predator, thus bestowing at least slight survival value on whatever factor was responsible for the delay.

Something akin to communicative displacement is involved in lobbing a stick or a stone around—it is like talking today about what one should do tomorrow. Of course it is not to be supposed that the first tool-carrying was purposeful, any more than that the first displaced communication was a discussion of phases. Caught in a cul-de-sac by a predator, however, the early hominid might strike out in terror with his stick or stone and by chance disable or drive off his enemy. In other words, the first tool-carrying had a consequence but not a purpose. Because the outcome was fortunate, it tended to reinforce whatever factor, genetic or traditional, prompted the behavior and made the outcome possible. In the end such events do lead to purposing behavior.

Although elements of displacement might arise in this fashion on the whole it seems likely that some degree of productivity preceded any great proliferation of communicative displacement as well as any significant capacity for traditional transmission. A productive system requires the young to catch on to the ways in which whole signals are built out of smaller meaningful elements, some of which may never occur as whole signals in isolation. The young can do this only in the way that humans learn their language: by learning some utterances as whole units, at the same time testing various blends based on that repertory, and finally adjusting their patterns of blending until the bulk of what they say matches what adults would say and is therefore understood. Part of this learning process is bound to take place away from the precise situations for which the responses are basically appropriate, and this means the promotion of displacement. Learning and teaching, moreover, call on any capacities for traditional transmission that the band may have. Insofar as the communicative system itself has survival value, all this behavior shows survival value also on the capacity...
for traditional transmission and for displacement. But these in turn increase the survival value of the communicative system. A child can be taught how to avoid certain dangers before he actually encounters them.

These developments are also necessarily related to the appearance of large and convoluted brains, which are better storage units for the conventions of a complex communicative system and for other traditionally transmitted skills and practices. Hence the adaptive value of the behavior serves to select genetically for the change in structure. A lengthened period of childhood helplessness is also a longer period of plasticity for learning. There is therefore selection for prolonged childhood and, with it, later maturity and longer life. With more for the young to learn, and with male as well as female tasks to be taught, fathers become more domesticated. The increase of displacement promotes retention and foresight; a male can protect his mate and guard her jealously from other males even when he does not at the moment hunger for her.

There is excellent reason to believe that duality of patterning was the last property to be developed, because one can find little if any reason why a communicative system should have this property unless it is highly complicated. If a vocal-auditory system comes to have a larger and larger number of distinct meaningful elements, those elements inevitably come to be more and more similar to one another in sound. There is a practical limit, for any species or any machine, to the number of distinct stimuli that can be discriminated, especially when the discriminations typically have to be made in noisy conditions. Suppose that Samuel F. B. Morse, in devising his telegraph code, had proposed a signal 1 second long for "A," 2 second long for "B," and so on up to 2.6 seconds for "Z." Operators would have enormous difficulty learning and using any such system. What Morse actually did was to incorporate the principle of duality of patterning. The telegraph operator has to learn to discriminate, in the first instance, only two lengths of pulse and about three lengths of pause. Each letter is coded into a different arrangement of these elementary meaningless units. The arrangements are easily kept apart because the few meaningless units are plainly distinguishable.

The analogy explains why it was advantageous for the forerunner of language, as it was becoming increasingly complex, to acquire duality of patterning. However it occurred, this was a major breakthrough; without it language could not possibly have achieved the efficiency and flexibility it has.

One of the basic principles of evolutionary theory holds that the initial survival value of any innovation is conservative in that it makes possible the maintenance of a largely traditional way of life in the face of changed circumstances. There was nothing in the make-up of the protohominoids that destined their descendants to become human. Some of them, indeed, did not. They made their way to ecological niches where food was plentiful and predators sufficiently avoidable, and where the development of primitive varieties of language and culture would have bestowed no advantage. They survive still, with various sorts of specialization, as the gibbons and the great apes.

Man's own remote ancestors, then, must have come to live in circumstances where a slightly more flexible system of communication, the incipient carrying and shaping of tools, and a slight increase in the capacity for traditional transmission made just the difference between surviving—largely, be it noted, by the good old protohominoid way of life—and dying out. There are various possibilities. If predators became more numerous and dangerous, any nonce use of a tool as a weapon, any co-operative mode of escape or attack might restore the balance. If food became scarcer, any technique for cracking harder nuts, for foraging over a wider territory, for sharing food so gathered or storing it when it was plentiful might promote survival of the band. Only after a very long period of such small adjustments to tiny changes of living conditions could the factors involved—incipient language, incipient tool-carrying and toolmaking, incipient culture—have started leading the way to a new pattern of life, of the kind called human.

SUBHUMAN PRIMATE CALLS are represented here by sound spectrograms of the roar (top) and bark (bottom) of the bowler monkey. Frequencies are shown vertically; time, horizontally. Roaring, the most prominent bowler vocalization, regulates interactions and movements of groups of monkeys, and has both defensive and offensive functions. Barking has similar meanings but occurs when the monkeys are not quite so excited. Spectrograms were produced at Bell Telephone Laboratories from recordings made by Charles Southwick of the University of Southern Ohio during an expedition to Barro Colorado Island in the Canal Zone. The expedition was directed by C. R. Carpenter of Pennsylvania State University.