SCIENCE

VOL. 98 FRIDAY, JULY 2, 1943 No. 2531

SCIENCE: A Weekly Journal devoted to the Advancement of Science, edited by J. MCKEEN CATTELL and published every Friday by

THE SCIENCE PRESS

Lancaster, Pennsylvania

Annual Subscription, $6.00 Single Copies, 15 Cts.

SCIENCE is the official organ of the American Association for the Advancement of Science. Information regarding membership in the Association may be secured from the office of the permanent secretary in the Smithsonian Institution Building, Washington, D. C.

THE ORIGIN OF LANGUAGE[1]

By Dr. E. L. THORNDIKE

PROFESSOR EMERITUS, TEACHERS COLLEGE, COLUMBIA UNIVERSITY

NOBODY knows when, where or how speech originated, and I am stepping in where wise scholars in linguistics and psychology fear to tread. My colleagues in psychology will, I beg, permit this divagation into speculation by one who has labored long in the less exciting fields of experiment and statistics. I ask and expect no mercy from experts in linguistic science, but only that they build a better theory on the ruins they make of mine.

We must first glance at three time-honored and then dishonored theories, now known by these opprobrious names: ding-dong theory, bow-wow theory and pooh-pooh theory.

The ding-dong theory assumed a mystical power of certain things to evoke certain sounds from men. Since each such sound was associated with the experience of the thing, it came to mean it. And since men were alike in their responses to things by sounds, one of these sounds meant more or less the same thing to all in the group, and easily became a vehicle of communication. All the evidence is against the existence of any such mystical power, and only extremely strong evidence would induce any scientific student of psychology or of language to put any faith in so extremely unlikely an origin of language.

The bow-wow theory supposed that men formed habits of using the sounds made by animals, things or events to mean the respective animals, things and events and that these habits started them on the road to inventing other sounds as signs of animals, things or events. For various reasons this theory is discredited. Doubtless after man has language, he will

[1] A lecture given on November 5, 1942, as one of the series of the William James lectureship at Harvard University.

often make the sounds that animals and things make, but it is doubtful how often he will do so in a languageless group. Possibly he will do so only accidentally as a part of his general vocal play. There might be little agreement in the ideas evoked in the members of a human group by hearing the varying sounds which its various members made when they thought of a dog, a cow, thunder and the like.

Even if a group got a sufficient agreement in the case of forty or fifty sounds for these to be used commonly in the group, an advance by the addition of non-mimetic sounds as signs of things and events would be difficult. If the mimetic sounds remained fully mimetic, it might well be impossible. But the opponents of the bow-wow theory have not considered sufficiently the possibility that a human group might modify their vocabulary of mimetic sounds by slurring, abbreviation and other processes that make speech easier for the speaker without losing the old meanings of animals, things and events in the hearer. If close imitations of a dog's barking, cock's crowing, baby's crying, lamb's bleating, etc., became conventionalized within a human group into sounds no more like the originals than *bow-wow, cockadoodledoo, mama* and *bah-bah,* or *urr-urr, uk a duk a duk duk, na-na* and *buh-buh,* that group could in a few generations progress to a set of sounds many of which would mean primarily certain animals or things and only secondarily or not at all the sound made by the respective animals and things. The group's vocabulary would all be about things that had distinctive sounds, but could be in the form of sounds different from these and in some cases hardly suggestive of them. The invention of a non-mimetic sound for some thing hitherto nameless would then be easier. The use of such an invention would, of course, spread somewhat slowly within the group and very slowly outside it to groups accustomed only to mimetic words.

The pooh-pooh theory, or interjectional theory, supposed that the instinctive unlearned cries of man as a wordless animal, which already are sounds that are evoked by certain situations and evoke in human hearers certain equally unlearned responses of action and feeling, came to possess meanings also, and that on the basis of this vocabulary of familiar sounds meaning pain, surprise, fear, affection and the like, early man here and there used other sounds to mean other facts.

Nobody should doubt that part of this is true. To a mother whose baby cries and seeks her breast that cry probably means that the baby wants to be fed if anything means anything to her. If she can think of anything she will think of that, as well as react appropriately to it. But for various reasons students of language have decided that the attachment of meanings to the hearing or the making of these sounds of instinctive nature is not adequate to originate articulate speech. So-called animal language plus the power of thinking meanings would not produce human language.

An ingenious theory has been set forth by Sir Richard Paget, a physicist and student of phonetics, who argues that the total behavior of a man to a situation includes characteristic movements of the tongue and lips and other organs of speech. These gestures of the mouth parts became specially important when a man's hands were "in continual use . . . for craftsmanship, the chase, and the beginnings of art and literature," so that he could not gesture with them. Sounds were added to these "mouthings," and finally came to play the leading role. In Paget's own words:

> Originally man expressed his ideas by gesture, but as he gesticulated with his hands, his tongue, lips and jaw unconsciously followed suit in a ridiculous fashion, "understudying" (as Sir Henry Hadow aptly suggested to me) the action of the hands. The consequence was that when, owing to pressure of other business, the principal actors (the hands) retired from the stage—as much as principal actors ever do—their understudies—the tongue, lips and jaw—were already proficient in the pantomimic art.
>
> Then the great discovery was made that if while making a gesture with the tongue and lips, air was blown through the oral or nasal cavities, the gesture became audible as a whispered speech sound. If, while pantomiming with tongue, lips and jaw our ancestors sang, roared or grunted —in order to draw attention to what they were doing—a still louder and more remarkable effect was produced, namely, what we call voiced speech. . . .
>
> In this way there was developed a new system of conventional gesture of the organs of articulation from which, as I suggest, nearly all human speech took its origin. . . .
>
> We can now form a mental picture of how the process of speech-making actually began, but an example or two will make the argument clearer. If the mouth, tongue and lips be moved as in eating, this constitutes a gesture sign meaning "eat"; if, while making this sign, we blow air through the vocal cavities, we automatically produce the whispered sounds mnyΔm–mnyΔm (mnyum), or mnIΔ–mnIΔ (mnyuh)—words which probably would be almost universally understood, and which actually occur as a children's word for food in Russian, as well as in English. . . .
>
> Another adult example may be given, namely, in connection with the beckoning gesture—commonly made by extending the hand, palm up, drawing it inwards towards the face and at the same time bending the fingers inwards towards the palm. This gesture may be imitated with the tongue, by protruding, withdrawing, and bending up its tip as it re-enters the mouth and falls to rest.
>
> If this "gesture" be blown or voiced, we get a resultant whispered or phonated word, like eda, eda or edra (according to the degree of contact between tongue and upper lip or palate) suggestive of the Icelandic hadr, the Hindustani idhar and the Slavonic ɪdeɪ—all of which bear much the same meaning as our English word

"hither." If the same tongue gesture be finished more vigorously, the resultant word will end in a *k* or *g*, owing to the back portion of the tongue making a closure against the soft palate.

Thus, by unconsciously using the tongue, lips, jaw, etc., in the place of the head, hands, etc., pantomimic gesture would almost automatically produce human speech.[2]

Paget fabricated words by moving his own jaws, tongue and lips in ways which seemed to him likely to have been used as oral gestures of primitive men accompanying manual or other gestures meaning reach up, draw back suddenly, scrape, wave aloft, shoot with a bow and arrow, sew, blow, plough, strip grains from the stalk, pick berries, collect them and bury them in the ground, and many others. He finds substantial correspondences between his fabricated sounds and certain words in old languages. Of the famous Aryan roots he considers that 77 per cent. are clearly pantomimic. For example, "tank—contract, compress—as in thong, is due to two compressions in succession fore and aft the palate." "da—give—seems to be an offering gesture made with the tongue."[3]

Paget's book "Human Speech" is so recent (1930) that his theory has not yet received a pet name. Using the first illustration that he gives we might call it the yum-yum theory. This, however, really misrepresents and unduly favors it; for the theory requires the mouth parts to pantomime not eating, drinking, sipping, blowing and other acts of the mouth parts themselves (nobody doubts that), but movements of other parts of the body. A truer nickname would be the "tongue-tied" theory, meaning that the tongue is yoked with the body by subtle bonds of mimetic kinship. The theory has been accepted by at least one psychologist, Eisenson, but it has not been acceptable generally. Personally, I do not believe that any human being before Sir Richard Paget ever made any considerable number of gestures with his mouth parts in sympathetic pantomine with gestures of his hands, arms and legs, still less that any considerable number of men in any local community made the same oral gestures in such pantomine.

And now for my theory, which is a humdrum affair compared with any of these four.

Let us assume a group of one or more human families living together at least as continuously as one of the groups of chimpanzees studied by Nissen in their natural habitat. Let us assume that their environment includes, besides the untouched objects of nature, a few objects chosen and preserved as tools, say a few pounders, a few cutters, a few gourds, shells or other dippers and holders, and perhaps a few stabbers and scrapers; and also some natural objects chosen and preserved as playthings, things that one can chew on, roll or throw, make a noise with, and the like.

We may safely assume further that these humans made a wide variety of movements with their hands, much the same as the human infants of to-day instinctively make, pushing, pulling, tearing, putting into their mouths, dropping out therefrom, dropping, throwing, picking up, etc., etc.

We may safely assume further that these humans made a variety of sounds like the meaningless prattle of infants, letting their mouth parts play with their voices in the same multifarious way that their hands play with any obtainable object. The variety of sounds made may indeed have been greater than that made by an infant of to-day, whose vocal play may be narrowed by the elimination of sounds which are alien to the language which his environment favors. And we know that an infant of to-day makes a much wider variety of articulate sounds than the language of his parents contains.

Such a person in such a group would at an early age have a memory image or expectation or idea of the appearance of the person who nursed him, which her voice or smell or caress could evoke though she was unseen. By having been experienced in so many different contexts, some image or expectation or idea referring to her would have acquired an existence independent of any particular sequence of behavior. In a similar manner he would have an image or expectation or idea referring to each object that had been associated with many varying concomitants in his uses of it or play with it.

Such a person would prattle while he worked or played much as a child of a year or two now prattles as he plays. If his making a certain sound became connected with his experiencing a certain object or act and having an image or expectation or idea of that object or act, he would have a language. That sound and the act of making it would mean that object or act to him. It would be a private language useless as yet for communication. It would be a narrow language consisting of only a few words referring mostly to his own acts and possessions, to the persons in the family group and to their acts and possessions. But it would be genuine language.

And it would be a valuable intellectual tool for its possessor, enabling him to replace the somewhat cumbrous and elusive images or expectations by sounds that he could make and arrange more or less at will. If he did connect *ik* with his digging stick, and *üg* with his large turtle-shell container, *yum* with truffle and *kuz* with clam, he could plan an expedition to get truffles or one to get clams more easily and conveniently than he could with only pictorial memories. Consequently, we may safely reckon that any person

[2] R. Paget, "Human Speech," pp. 133–138, *passim*, 1930.

[3] *Ibid.*, p. 149.

who made these connections that gave sounds meanings and gave things symbolic equivalents would keep them, even though he alone understood them.

What now is the probability that a person brought up in a languageless family group would form one such connection whereby a sound (not an instinctive cry of pain, delight, triumph, etc.) meant an object? What is the probability that he would form two such? Three? Four? A dozen? A score?

Properly planned experiments with enough infants brought up in a languageless environment for ten years (perhaps for a much shorter time) would give a decisive answer. I have long wished to make systematic observations of infants in linguistically underprivileged environments, but have never been able to find the time, and must rely upon memories of casual observations of my children and grandchildren in making my estimates.

I think that the probability that a person in the top half of the species for intelligence by birth would make four or five such connections is very high, say seven out of ten. Consider a child of early man playing with a large shell used as a container in the household and prattling as he plays. Let us take the state of affairs least favorable to connecting the sound *üg* with that shell.

Let his prattling possibilities consist of a thousand syllables all equally likely to occur, and all as likely to occur in any one situation as in any other. Then the chance that he will utter *üg* as he puts a pebble in the shell is 1 in 1,000 if he prattles at all. And unless that connection between the manual act and the vocal act is somehow strengthened, he will be as likely the next time that he drops a pebble into that shell to utter any other sound in his repertory as to utter *üg*. Very often he will utter other sounds and no progress will be made toward the attachment of meanings to his utterances.

But there are forces which tend to cause progress away from purely miscellaneous vocal play. First of all the child who puts one pebble in the shell is likely to put another in then and there. His enjoyment of the act makes him repeat it, that is, strengthens its connection with the mental set in which he did it first. Now that mental set happened at that time to evoke also the vocal play of saying *üg*, and the confirming reaction which the enjoyment of the manual play set in action tends to spread or scatter so as to strengthen also the connection of the situation with the utterance.

In the second place, saying *üg* to the shell and pebble may be itself enjoyable and the connection may thereby be strengthened. Consequently, the probability that the child will drop a second pebble is substantial and the probability that he will utter *üg* therewith if he utters anything is far above 1 in 1,000.

Let us assume provisionally that some active-minded Homo Sapiens did thus connect *ma* with the mother who nursed and fondled him, *ba* with the round black thing that rolled and tossed, *unk* with the club with which he knocked down his prey, and similarly for a dozen or more "words," as we may truly call them. If he did this, what would be the probability that some second person in the group would come to understand these words? And if he did come to understand them, what would be the probability that the first person would come to use them with the intent of having the second person understand them, and so attain the condition of possessing speech as a social tool?

If one person in a hitherto languageless group of two or three dozen souls has reached the stage of a private language of a score of words the probability that some other person in the group will come to understand three or four of his words is much more than infinitesimal.

His companions might well hear him say *kuz* as he dug up a clam or opened a clam or ate a clam, a hundred times in a week. Even if they paid no more attention to his speech than to his personal play, vocal or non-vocal, the sound *kuz* would tend to make them think of a clam more often than of any other one object. And under certain conditions they would be attentive to his speech. For example, in a group digging for clams together, if one cried *kuz* whenever he found a clam, the cry would become interesting to others.

If the group had a dozen or so "bow-wow" (that is, mimetic) words that they used as signals, they would be thereby the more disposed to attend each to the other's vocalizations. If a second person of the group had a private language of his own, though unlike that of the first person in every particular, the second person would be thereby the more disposed to attend to the first person's vocalizations. If the group had a system of mutual influence by gestures, even one utterly devoid of any vocal accompaniments, its members would be thereby a little more disposed to attend to the vocal behavior one of another.

So I would set the probability that in a group of thirty souls, one of whom had a private language of twenty words, some one other person would come to understand five of these words in the course of a moderate lifetime of thirty-five years as well above one in ten thousand, and probably above one in a thousand.

If the family group of say thirty souls has an inventor of a private language of say twenty-five words and say ten of the thirty understand say eight of the words, what is the probability that any one of these ten will use any of the eight words that he understands, use it, that is, to mean to himself the thing or act or event in question? This probability is substantial, but it is not 100 out of 100. Some persons

in such a group will hear and understand a word hundreds of times, but in all probability never say it at all, except accidentally as an element in their meaningless chatter. But some will, when they themselves utter this word in their meaningless chatter or for any reason, understand it as if it were spoken by A. And this act of saying a word and having it mean something will tend to be satisfying rather than annoying. Meaningful prattle is more satisfying than meaningless and will therefore be more frequently repeated.

If A, the original inventor, hears B or C or D say one of the words to which he, A, attaches meaning when he himself says them, what is the probability that he, A, will understand the word spoken by B? It is not 100 out of 100. The connection *kuz*→*clam* may remain confined to *kuz* said by A, because A is stupid or by nature an extreme introvert or what William James called a lonely thinker, or because of the general tendency of connections to operate only in the way in which they are formed. But A has, by hypothesis, an IQ of 100 or better, and if B goes about saying *kuz* repeatedly and as if he meant something, A is likely to notice what B says, and will at least be more likely to attach the thought of clam to the sound *kuz* when made by B than to attach any other one meaning to it. I should conjecture that the probability of A's understanding B would be well over 25 in 100 and under 90 in 100.

It is perhaps time to attach a name to the theory which I am expounding. Let us save everybody trouble by giving it an opprobrious name from the start! Since it relies on the miscellaneous vocal play of man instead of his alleged mimetic or emotional utterances, it could be called the "babble-babble" theory. Since it starts with languages private to single persons, and progresses gradually toward speech in the full speaker-hearer relation (which, indeed, my exposition has not yet reached) it could be called the "onety-twoty" theory. Since it depends on successive selections of chance variations in sound-reality connections, it could be called the "chancey-chance" or "luck-luck" theory. Or we may combine its two main dynamic features and call it the babble-luck theory.

Let us continue with the luck-luck course of the babble-luck theory.

If B understands *kuz* as spoken used by A and A understands *kuz* as spoken by B, what is the probability that A will come to use the word as a means of influencing B? What is the probability that B will come to *use* the word to influence A?

It is not 100 out of 100. A and B might continue for years to get meaning from one another's use of the word, but never use it for any purpose other than as a self-reminder or as an aid in personal plans or for self-entertainment. However, if A said *kuz* when he was about to set forth to dig clams, and B was moved by hearing *kuz* to set forth to dig clams also, and so accompanied A on several occasions, there might fairly easily be built up a habit in A of saying *kuz* when he wanted B's company on a clamming trip. (The formation of this habit would not be as simple as this sounds or by one direct linking, but by various cooperating associative links which I could describe if necessary). Or if A had already a habit of purposive communication with B by means of a gesture such as pointing to a clam and to B's mouth when he wished or permitted B to eat it, A might well happen to say *kuz* along with the two gestures and eventually in place of the former gesture. (Here again the substitution would not be as simple as it sounds, but it could come to pass.)

A and B thus reach a stage where a word is used by one of them, say by A, with the expectation that his saying it in the presence of the other will produce or favor certain behavior in the other, and where A has the habit of saying it to the other as an appropriate thing to do when a certain desire or purpose moves him. This is genuine human language used in the speaker-hearer relation. But the relation is, as yet, unidirectional, from A as speaker to B as hearer.

Speech need not progress further to full two-way, give-and-take speech, but it could, and often would. I will not run the risk of wearying you with the probabilities that the normal operations of repetition and reward would lead men to this final stage. They are high.

Each of the stages that I have described, from that of words used privately to purposive use of speech in the full speaker-hearer relation, was self-sustaining, by adding something to the group's balance of satisfactions, or to its chance of survival, or to both. A one-man language could make that man remember, anticipate and plan better. In so far as others understood A's words, each of them had some profit from A's experience in addition to their own. In so far as they used his words, each had a private language without originating it. When they reached the stage of understanding one another certain experiences of any one were of profit to all. The stage of purposive use of words to modify the behavior of another gave the possibility of increasing costless cooperation and decreasing costly interference of person with person. Even if the words used were few and the occasions of their use limited to a very narrow round of suggestions, commands, invitations and reports, the benefits would still be enough to maintain the linguistic activities.

Nothing in all this so far requires that either A or B thinks of the other as imagining or meaning clams when he says *kuz*. Such imputation of an inner life to another may arise later and regardless of communication, though of course it can not progress far

without communication. How it arises is a fascinating problem, but to discuss it would make far too long an interruption of our present task.

Let us turn rather to some possible criticisms. First it will be said that the speech which I have derived from babble by luck is a pitifully small, crude affair in comparison with the speech of any known group present or past. This criticism is true. Even after a dozen or more words had been used purposively hundreds of times by a third of the family group and understood after a fashion by two thirds of the group, the use and understanding would be nowhere nearly as clean-cut as that of a modern man or child. A person could use words more or less appropriately in certain situations in the sense that the use of the word was much more appropriate on the average than saying nothing, or than saying some other word of those in his active vocabulary. He could understand words in the sense that what he did to the total situation including the word was on the average different from what he would have done if some other word had been there, and better than what he would have done if no word had been there. But when the imperfect appropriateness of a speaker's uses was combined with a hearer's inadequate understandings, a perfect result could not be expected. If the speaker went much beyond the regular routine uses, he would arouse misunderstanding, neglect or perplexity. The group's linguistic activities might be clumsy as well as extremely narrow.

It will be said that the evolution of any language worthy of the name from such crude beginnings is problematic. This criticism also is true, but it is not very damaging. The problems are no harder than the problems of the evolution of mechanical tools from their crude beginnings. The evolution of a vocabulary of two hundred names of acts, objects and events from a vocabulary of twenty is a problem, though a rather easy one.[4] The evolution of a language that can mean qualities and relations as well as objects and events is a further problem. The evolution of a language that can by sounds ask questions, distinguish orders from statements and date events has further problems. Refinements of meaning, as by our adjectives and adverbs, and abbreviations of speech, as by our pronouns, involve further problems.

I have not solved these and other problems. But I think they are all soluble. If the facts which I have related account for how men came to use articulate words with the purpose of influencing other men, to understand such words and to cooperate in the speaker-hearer relation, they can fairly be said to account for the origin of language, but to leave us with many problems of its development.

A third possible criticism is that the babble-luck doctrine should have produced dozens, maybe hundreds of different languages of this beggarly sort. Origin from miscellaneous babble would cause a multiplicity of primeval languages unless one family group got so great a head start that its language spread to all other tribes before they had invented any languages of their own, which is unlikely. I see nothing objectionable in this. It seems to me sure that any continuing group of intelligent human beings would in time get a language from "babble and luck" if they did not get it earlier from neighbors or visitors who already had it. In many cases they would get it so. Inter-group learning would be of the same general nature as the intra-group learning.

A fourth possible criticism is that hundreds of generations seem to be required to get even this beggarly language if the group has no aid from outside. This seems to be really an argument pro rather than con. Surely the notion that primeval men who were wordless got words as quickly as modern men got Mohammedanism or Christianity or steam engines is fantastic. The length of time from selecting and using flints that were sharp to chipping flints to make them sharp, and the length of time from chipping them roughly to chipping and polishing them in the elegant neolithic styles are both reckoned in many thousands of years.

Whatever may be the value of this account of the origin of meaningful speech, one thing is certain. The human animal's miscellaneous play with his vocal apparatus and the articulate sounds he thereby produces and the associations he makes of these with things and events independently of, and especially contrary to, his linguistic environment deserve much more attention from psychology and linguistic science than they have hitherto received.

OBITUARY

CHARLES FREDERICK MARVIN

DR. CHARLES FREDERICK MARVIN, former chief of the United States Weather Bureau, died in the early morning of June 5, 1943, at Doctors Hospital, Washington, D. C., of heart failure following a recent operation. He was born at Putnam (now Zanesville), Ohio, October 7, 1858, son of Charles F. and Sarah A. (Speck) Marvin; and was educated at the public

[4] One generation having reached the linguistic status I have described, the second generation can learn from it and spend most of its linguistic activity in adding its inventions to the parental stock. The custom of naming things and acts by sounds may, after a certain number of such sound-meaning connections has been reached, become a conscious deliberate habit. Some early linguist may then devote his spare time to naming every person in the group, every animal that frequents the locality, and every tool or weapon that he uses.