Knowledge Unlearned and Untaught: What Speakers Know about the Sounds of their Language

Morris Halle 1978

Untaught Knowledge

The native speaker of a language knows a great deal about his language that he was never taught. An example of this untaught knowledge is illustrated in (1), where I have listed a number of words chosen from different languages, including English. In order to make this a fair test, the English words in the list are words that are unlikely to be familiar to the general public, including most crossword-puzzle fans:

(1) ptak thole hlad plast sram mgla vlas flitch dnom rtut

If one were to ask which of the ten words in this list are to be found in the unabridged Webster's, it is likely that readers of these lines would guess that *thole*, *plast*, and *flitch* are English words, whereas the rest are not English. This evidently gives rise to the question: How does a reader who has never seen any of the words on the list know that some are English and others are not? The answer is that the words judged not English have letter sequences not found in English. This implies that in learning the words of English the normal speaker acquires knowledge about the structure of the words. The curious thing about this knowledge is that it is acquired although it is never taught, for English-speaking parents do not normally draw their children's attention to the fact that consonant sequences that begin English words are subject to certain restrictions that exclude words such as *ptak*, *sram*, and *rtut*, but allow *thole*, *flitch*, and *plast*. Nonetheless, in the absence of any overt teaching, speakers somehow acquire this knowledge.

The sounds of speech In order to get some insight into how humans acquire knowledge about their language without being taught, it is necessary to understand the character of the knowledge that is being acquired. Since I am talking about sounds and sound sequences, I must

Reprinted with permission from M. Halle, J. Bresnan, and G. A. Miller (eds.) *Linguistic Theory and Psychological Reality*. MIT Press, Cambridge, MA, pp. 294–303, 1978.

say a few words about the way that linguists think about the sounds of speech. These ways of thinking about sounds derive in part from the work of Alexander Graham Bell and that of his father, A. Melville Bell. Let us turn, therefore, to the Bells' contribution to the science of language.

Alexander Graham Bell was a speech therapist by profession: his specialty was the teaching of speech to the deaf, and according to all reports he was an extraordinarily gifted and successful practitioner of this difficult art. Speech therapy was the profession of many members of the Bell family. In fact, it was a sort of family enterprise. The head of the family, A. Melville Bell, practiced it in London; other members, in other parts of Great Britain. What differentiated A. Melville Bell from most speech therapists was that he was interested not only in the practical aspects of his work, but also in its scientific foundations. In this work he involved his son, the future inventor of the telephone, and on one issue of importance the son made a contribution that went far beyond that of his father.

A. Melville Bell's analysis of spoken language proceeds from the obvious observation that the production of speech sounds involves the coordinated activity of a number of different organs such as the lips, the tongue, the velum, and the larynx, which together make up what traditionally has been called the human vocal tract. From this point of view the act of speaking is an elaborate gymnastics or choreography executed by different speech organs. In A. Melville Bell's book *Visible Speech* (1867) we find a systematic account of the different activities that each speech organ is capable of, together with a discussion of the different speech sounds that result from particular combinations of activities of specific speech organs.

Consider from this point of view the initial consonants in the words *veal*, *zeal*, *sheep*, *keel*, *wheel*. One thing that differentiates each of these consonants from the others is the place in the vocal tract that is maximally narrowed and the organ or organs effecting this narrowing. In /v/ the constriction is formed by raising the lower lip; such sounds are therefore designated as *labial*. In /z/ and /š/ the constriction is formed with the tongue blade, and these sounds are designated by the term *coronal*. In /k/ the constriction is formed with the dorsum (or body) of the tongue and such sounds are designated as *dorsal*. The sound beginning the English word *wheel* is produced with two simultaneous constrictions, one with the lips and the other with the tongue dorsum; this sound is therefore both *labial* and *dorsal*.

A further mechanism involved in distinguishing one sound from another is voicing—whether or not the sound is produced with the accompaniment of vibration of the vocal cords: /z v/ are; /š k x^w/ are not. This fact can readily be verified by placing one's finger tips on the large (thyroid) cartilage in the front of the neck and pronouncing the sounds in question. When the vocal cords vibrate, one can detect a slight throbbing sensation in the finger tips. Finally, for purposes of this discussion one additional mechanism must be identified. It is the mechanism that produces strident sounds, such as /f v s z š ž č ǰ/, and distinguishes them from the rest. It consists in directing the air stream against the sharp edges of the upper teeth, thereby producing audible turbulence.

Thus five distinct mechanisms that are involved in the production of the continuant sounds under discussion have been identified. I label these for present purposes as follows:

the raising of the lower lip—labial
the raising of the tongue blade—coronal
the raising of the tongue body—dorsal
vocal cord vibration—voicing
air stream directed at upper teeth—strident

When two or more mechanisms are activated, the perceptual effect is that of a single sound. Thus, both /z/ as in *zeal* and /s/ as in *seal* are perceived as single sounds, although in the production of /z/ one more mechanism (voicing) is activated than in the production of /s/. As shown in Figure 1, Bell's *Visible Speech* alphabet had a special symbol to represent each of these mechanisms; for example, the labial mechanism is represented by a semicircle open to the left, the coronal mechanism by a semicircle open to the top, voicing is symbolized by a line inside the semicircle, and so forth. When two or more mechanisms are activated simultaneously in the production of a given sound the symbolic representation becomes rather cumbersome. It is therefore more convenient to represent the same information by means of a matrix such as that in Table 8.1.

The claim made explicitly by A. Melville Bell in *Visible Speech* is that he had identified all mechanisms that are relevant in the production of sounds in any spoken language. If this claim is correct, it should be possible for an appropriately trained person to analyze any sound whatever in terms of the mechanisms involved in its production, espe-

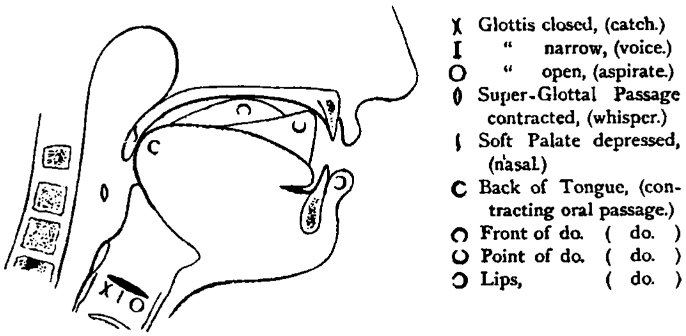

Fig. 1. Diagram of the Human Vocal Tract. The symbols on the right refer to the letters of Bell's phonetic alphabet. Reproduced from A. M. Bell, *Visible Speech* (1867).

cially since the number of mechanisms is fairly small. Moreover, it should also be possible for a trained person to produce sounds represented in this notation, even sounds that he had never heard before. That is exactly how Bell saw the matter and he set about demonstrating it in a most dramatic fashion. The following description of a demonstration is from a letter written by an observer, Alexander J. Ellis, Esq., F.R.S., which Bell quotes in *Visible Speech*.

Table 8.1

		Labial	Coronal	Dorsal	Voiced	Strident
f	feel	+	–	–	–	+
v	veal	+	–	–	+	+
x^w	wheel	+	–	+	–	–
s	seal	–	+	–	–	+
z	zeal	–	+	–	+	+
š	sheep	–	+	–	–	+
ž	rouge	–	+	–	+	+
č	cheap	–	+	–	–	+
ǰ	jeep	–	+	–	+	+
x	Bach	–	–	+	–	–
p	peal	+	–	–	–	–
d	deal	–	+	–	+	–
k	keel	–	–	+	–	–

The mode of procedure was as follows: Mr. Bell sent his two Sons, who were to read the writing, out of the room—it is interesting to know that the elder, who read all the words in this case, had only five weeks' instruction in the use of the Alphabet—and I dictated slowly and distinctly the sounds which I wished to be written. These consisted of a few words in Latin, pronounced first as at Eton, then as in Italy, and then according to some theoretical notions of how Latins might have uttered them. Then came some English provincialisms and affected pronunciations; the words 'how odd' being given in several distinct ways. Suddenly German provincialisms were introduced. Then discriminations of sounds often confused.... Some Arabic, some Cockney-English, with an introduced Arabic guttural, some mispronounced Spanish, and a variety of vowels and diphthongs.... The result was perfectly satisfactory;—that is, Mr. Bell wrote down my queer and purposely-exaggerated pronunciations and mispronunciations, and delicate distinctions, in such a manner that his Sons, not having heard them, so uttered them as to surprise me by the extremely correct echo of my own voice.... Accent, tone, drawl, brevity, indistinctness, were all reproduced with surprising accuracy. Being on the watch, I could, as it were, trace the alphabet in the lips of the readers. I think, then, that Mr. Bell is justified in the somewhat bold title which he has assumed for his mode of writing—"Visible Speech." (p. 22)

The quaintness of this testimonial should not be permitted to obscure the serious point that Bell attempted to establish by means of his demonstration, namely, that all sounds of all known languages can be produced, given the very restricted information about a small number of mechanisms that is provided by *Visible Speech*. Anybody who controls all the mechanisms singly and in combination can produce any speech sound whatever. It is therefore these mechanisms and not the individual sounds of language that are the fundamental building blocks of speech. This insight, which in the last quarter century has become almost a truism among students of language, was stated explicitly in the early 1900s by Alexander Graham Bell in a series of lectures that he delivered to the American Association to Promote the Teaching of Speech to the Deaf. (It should be noted that Bell's terms "constriction" and "position" are synonymous with what has been termed "mechanism" here.)

What we term an "element of speech" may in reality ... be a combination of positions. The true element of articulation, I think, is a constriction or position of the vocal organs rather than a sound. Combinations of positions yield new sounds, just as combinations of chemical elements yield new substances. Water is a substance of very different character from either of the gases of which it is formed; and the vowel *oo* is a sound of very different character from that of any of its elementary positions.

When we symbolize positions, the organic relations of speech sounds to one another can be shown by means of an equation; for example

English wh = P (labiality—MH) + P′ (dorsality—MH)
German ch = P′ (dorsality—MH)
hence German ch = English wh − P (labiality—MH)

The equation asserts that the English *wh* without labial constriction is the German *ch*. (*The Mechanism of Speech*, pp. 38–39)

Sounds into words I now turn from the analysis of speech sounds into their component mechanisms—or features, to use a more modern term—to the restrictions that languages characteristically impose on the concatenation of sounds into words. We have already seen in (1) that certain consonant sequences are not admissible at the beginning of English words. Hence the words beginning with the sequences *pt*, *hl*, *sr*, *mgl*, *vl*, *dn*, and *rt* were judged not to be part of the English lexicon. A different kind of restriction is found in the choice of the plural marker in English. I have listed in (2) three different sets of English nouns:

(2) a. bus, bush, batch, buzz, garage, badge
b. lip, pit, pick, cough, sixth
c. cab, lid, rogue, cove, scythe, cam, can, call, car, tie, gnu, blow, tray, sea, . . .

If you say to yourself the plural forms of the words in (2), you will notice that English has, not one, but three plural suffixes, one for each of the three separate sets of words in (2). We add an extra syllable /ɪz/ in forming the plural of the words in (2a); we add /s/ for the plural of the words in (2b), and we add /z/ to form the plural of the words in (2c). One can readily show that it is not the case that we memorize the plural form of every word we learn, for we know how to form the plurals of words we have never encountered before. Specifically, think of the plurals of the three English words in list (1): *flitch*, *plast*, and *thole*. I am sure that most readers who have never heard these words would agree that they know their plural forms and that these are respectively

flitches, like *buses* (2a)
plasts, like *lips* (2b)
tholes, like *cabs* (2c)

These facts show that speakers of English know a rule of plural formation. Like the restrictions on word-initial consonant sequences illustrated in (1) the English plural rule is rarely (if ever) overtly taught; many readers who have faithfully followed it all their lives may never have been aware of it until reading the preceding paragraph.

It is necessary to be clear about the status of a rule such as the plural rule under discussion. It is part of the knowledge that English speakers have and that people who do not know English normally do not have. Knowing the rule that determines the phonetic actualization of the plural in English is therefore much like knowing that the device invented by Alexander Graham Bell is called *telephone* rather than *farspeaker* (compare *loudspeaker*), *phonex*, or *glub*. The main difference between knowing the rule for the plural and knowing the word *telephone* is that the latter is conscious knowledge about which the speaker can answer direct questions, whereas knowledge of the plural rule and similar matters is largely unconscious and parts of it might conceivably never be accessible to consciousness. This fact, it should be noted at once, does not render such knowledge inaccessible to study by psychologists or linguists—that is, to scientists whose subject of inquiry is the speaker and his knowledge. Tacit knowledge can be established by the same methods that were used to establish other things inaccessible to direct observation, such as the nature of the chemical bond or the structure of the gene.

The question to be answered is in what form does the English speaker internalize his knowledge of the plural rule? An obvious candidate is (3):

(3) a. If the noun ends with /s z š ž č ǰ/, add /ɨz/;
b. Otherwise, if the noun ends with /p t k f θ/, add /s/;
c. Otherwise, add /z/.

It is important to note that this rule is formulated in terms of speech sounds rather than in terms of mechanisms or features. In the light of the above discussion, which suggested that features rather than sounds are the ultimate constituents of language, I shall now attempt to reformulate the rule in terms of features. The first move that one might make might be to replace each of the alphabetic symbols in (3) by its feature composition as shown in Table 8.1. Specifically, this means that one might replace /s/ by the feature complex [nonlabial, coronal, nondorsal, nonvoiced, strident]; /z/ by the same set of features except that in place of [nonvoiced] it would contain the feature [voiced]; and so on. It is not easy to see where such a translation of the rule into feature terminology gets us. In fact, it gets us nowhere until we observe that given a matrix like that in Table 8.1 it is possible to designate groups of sounds by mentioning one or two features. Thus, for example, if we asked for all and only sounds that are labial we would get the group /f v x^w p/, whereas if we asked for the sounds that are strident we get /f v s z š ž č ǰ/. Suppose now that we were to utilize this idea in the formulation of the plural rule and characterize each of the different lists of sounds by the minimum number of features that suffice to designate the group unambiguously. We should then get a rule much like (4) in place of (3).

(4) a. If the noun ends with a sound that is [coronal, strident], add /ɨz/;
b. Otherwise, if the noun ends with a sound that is [nonvoiced], add /s/;
c. Otherwise, add /z/.

Having formulated an alternative to the rule given above as (3), our task now is to determine which of the two alternatives is the one that English speakers use. The test we shall use is one suggested to me some years ago by Lise Menn. It consists of asking English speakers to form the plural of a foreign word that ends with a sound that does not occur in English. A good example, Ms. Menn suggested, is the German name *Bach* as in *Johann Sebastian*——, which ends in the sound symbolized by /x/. If English speakers were operating in accordance with rule (3), they would have to reject options (a) and (b) and form the plural in accordance with option (c); that is, they would say that the plural of /bax/ is /baxz/ with a word-final /z/. If, on the other hand, English speakers were operating in accordance with rule (4), they would have to perform a feature analysis of /x/ which would tell them that the sound is [nonlabial, noncoronal, dorsal, nonvoiced, nonstrident]. Given this feature composition, the plural of /bax/ could not be formed in accordance with option (a) since /x/ is neither [coronal] nor [strident]; it would have to be formed in accordance with option (b) since /x/ is [nonvoiced]. In other words, if speakers operated in conformity with rule (4), their output would be /baxs/, which, as is perfectly obvious, is also the response that the majority of English speakers would make. We must, therefore, conclude that the formulation (4) of the plural rule in terms of features, and not the formulation (3) in terms of speech sounds, correctly represents the knowledge of English speakers.

Unlearned Knowledge

There is yet another, more important, inference to be drawn from the fact that English speakers can apply the plural rule to a word ending with a sound that is not part of the repertory of English. In order to apply the rule, the speaker has to be able to establish that the foreign sound in question is nonvoiced. He must therefore have knowledge that allows him to determine the phonetic mechanism involved in the production of a sound that is not part of his language. The curious thing about such knowledge is that not only is there no indication that it might ever have been taught to speakers, there is also no indication that speakers could ever have acquired such knowledge. Think what evidence would have to be marshaled to support the claim that the knowledge in question was acquired. One would have to point to experiences in the life of the average English speaker that would permit him to acquire knowledge that is otherwise possessed only by phoneticians who have undergone rigorous training of the type Alexander Graham Bell received from his father. As this is obviously implausible, one is led to contemplate the possibility that at least some knowledge available to speakers is innate. In fact, there appears to be a certain amount of independent evidence that knowledge of the feature composition of sounds is available to children long before they could possibly have learned a language. Experiments conducted by Peter Eimas (1971) at Brown University have established that the ability to discriminate voiced from nonvoiced speech sounds is present in children practically at birth. The suggestion that the ability to determine the feature composition of speech sounds is innate has, therefore, a certain amount of experimental support.

This brings me to the end of what I have to say about the knowledge that speakers have of their language. What remains for me to do is to indicate how the information just reviewed helps us in trying to understand manifestations of the human cognitive capacity in domains other than language, how it might help us understand the human capacity to draw inferences, perform computations, play games with elaborate rules, interact with one another, and uncover significant truths about the nature of the world around us and within us. If these manifestations of man's mind are at all like language, then we must expect to find that large portions of the knowledge on which they are based will be inaccessible to consciousness, that some of this knowledge will be innate, and that only a modest fraction of the total will have been acquired as the result of overt teaching.

Note

This chapter is adapted from a paper presented at the Convocation on Communications, in celebration of the Centennial of the Telephone, MIT, March 9–10, 1976, published in *The Telephone's First Century—and Beyond* (Thomas Y. Crowell, New York, 1977).

References

Bell, Alexander Graham
1911 *The Mechanism of Speech.* New York and London: Funk & Wagnalls.
Bell, A. Melville
1867 *Visible Speech: The Science of Universal Alphabetics.* London: Simkin, Marshall; London and New York: N. Trubner.
Eimas, P., E. Siqueland, P. Jusczyk, and J. Vigorito
1971 Speech perception in infants. *Science* 171:303–306.