Colloques Internationaux du C.N.R.S.
N° 206. — PROBLÈMES ACTUELS EN PSYCHOLINGUISTIQUE

SOME IDEAS CONCERNING THE ACQUISITION OF PHONOLOGY *

John MORTON
Medical Research Council, Cambridge
and
Neil V. SMITH
School of Oriental and African Studies, London

It is well-established that the mapping of the acoustic signals of speech onto a phonemic representation is complex. Liberman, Cooper, Harris, MacNeilage, and Studdert-Kennedy (1967) and Liberman, Cooper, Shankweiler, and Studdert-Kennedy (1967), among others, have pointed out that the invariant features of a particular phoneme can only be expressed in terms of an abstract motor code. It is supposed, then, that the speech-analysis system develops by reference to the articulatory system. Other authors have suggested that " the constraints of speech production [might be] structured into the speech perception system " (Lieberman, 1968, p. 1584), or that " experience with the generation of speech movements and with the simultaneous observations of the acoustic consequences of these movements plays an important role in shaping the process whereby speech is perceived. " (Stevens, 1968, p. 102.)

THE DATA

Smith (1973) has analyzed the phonology of his own child, A, and has made a number of discoveries which bear on this problem. Throughout the period of intensive observations, from 2.2 to 4.0, A made systematic mispronunciations. It is important to note that none of these errors can be

---

* We are grateful to Byron Morgan for providing a program for the Wilcoxon tests and to Keith Tayler for his kind assistance with data processing.

attributed to perceptual problems. From the age of 2, A was able to make all distinctions between minimal pairs at all word positions except in unstressed syllables. As he developed, the nature of the mispronunciations changed. Table 1 gives examples of one set of confusions which occurred

TABLE 1. A sample set of *A*'s mispronunciations at 2:6.

queen	/ki:m/	squeeze	/ki :p/
quite	/kaip/	switch	/wit/
quack	/kap/	sweet	/wi:t/
twice	/taip/	clean	/ki:n/
twit	/tip/	scream	/ki:m/

when A was two and one-half. Smith has shown that this set of transformations can be accounted for most simply by the operation of a number of ordered rules. The rules mainly relevant for these particular examples may be roughly characterized as follows :

R1 Delete any preconsonantal /s/.

R2 For any string C_1/w/VC_2, where C_1 and C_2 represent any consonant and V any vowel, C_2 is made bilabial. Thus a final /n/ becomes /m/, /z/ becomes /b/, /t/ becomes /p/, etc. This rule must follow Rule 1 to account for the pronunciation of " switch " as " wit. " If the rules were the other way round, the word would become " wip " instead.

R3 If any of the continuents /w/, /l/, /y/ and /r/ occur immediately after another consonants, they are deleted.

R4 The characteristics of final consonants are altered such that only nasals and voiceless stops occur.

As the child developed, these rules were dropped out or modified as his speech came to approximate adult speech more nearly. In the course of the development Smith made the following observations.

1. The child developed his phonology in discontinuous steps indicating not that the pronunciation of single words had changed, but rather that a rule or set of rules had been changed.
2. Between successive stages, A would often alternate between the two rule systems in the pronunciation of a particular word, but apart from this his forms were completely consistent.
3. When his mispronunciations were repeated back to him, either by an adult or by means of a tape recording of his own voice, they were understood, albeit reluctantly. The child may indicate he had understood the word, but he also indicates that he knows the pronunciation is incorrect.

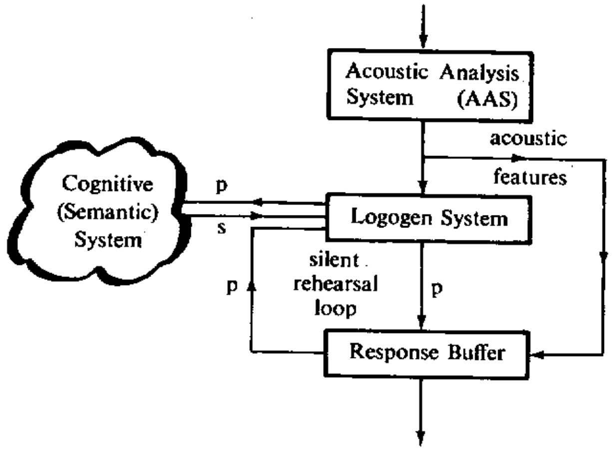
This is normal. Usually children react strongly, especially if they feel they are being made fun of.

4. If the repeated word was a homonym for the child, he would normally understand it as if it were the adult form. Thus, in response to " What is ' keep ' ? ", he would respond by collecting some objects and saying " These are mine. " On being pressed, however, and with some prompting he would, with considerable difficulty, retrieve the alternate meaning, in the case of " keep " indicating the adult " squeeze " by wringing a cloth. This difficulty is revealed in the following conversation which took place at a later time when A always pronounced as initial " sh " as " s. " Thus " shirt, " " shoe " and " ship " became " cert, " " soo " and " sip " respectively.
   N : What's a " cert ? "
   A : (points to his shirt)
   N : What's a " soo ? "
   A : (points to his shoe)
   N : What's a " sip ? "
   A : When you drink (and he imitates the action of drinking).
   N : What else does " sip " mean ?
   A : (puzzled and doubtfully says)... A zip ?
   (he distinguishes completely between /z/ and /s/ in his speech).
   N : No, it goes in the water.
   A : A boat !
   N : Say it.
   A : No, I can only say " sip ".
5. The most critical observation was that as soon as a rule had been changed the child could no longer provide any interpretation of the nonsense word which had just previously been part of his vocabulary. His typical response to " keem " (= queen) even to an earlier tape recording of his own voice (which he recognized) was " I don't understand what ' keem ' means. "

   A's inability to recognize his own earlier forms was maintained even when the meaning might easily be inferred from the context. At one stage in his speech he reduced an initial /st/ to /s/. A week or so after he had acquired a full initial /st/ he was listening to a tape recording of himself saying " Can I *sart* the tape recorder Daddy ? " He could not understand this sentence at all, asking " What was I wanting to do with the tape recorder ? "

THE MODEL

These observations do not seem to be explicable in terms of any model which suggests that the recognition of speech is directly mediated by the articulation system. We propose instead an account which is based on the Logogen Model (Morton, 1964, 1968*a*, 1969*a*, 1970). The relevant parts of this model are indicated in Figure 1.



p = phonological code
s = semantic code

FIG. 1. The elements of the Logogen System.

The central concept is that of a " Logogen " which has the function of making words available as responses. It is supposed that each word (or, more properly, morpheme) in the vocabulary is represented by a different logogen, a transducer between different kinds of coding. Each logogen is supposed to receive, from the other parts of the model, an indication of the presence of information which is relevant to the production or recognition of the word it represents. When more than a certain amount of information is present the logogen " fires, " sending a code to the Cognitive System or to the Response Buffer. In speech production the input to the Logogen System would be produced by the Cognitive System which would specify the semantic form of an intended word ; for word recognition the inputs would come from the Acoustic Analysis System ; in the recognition of continuous speech a combination of these kinds of signals would be used, accounting for the facilitative effects of context on word recognition. When the Logogen System operates to make a word available as a response, a code passes to the Response Buffer. Within the model, the Response Buffer has a number of functions. To start with it is necessary to have some means of distinguishing a response being available from a response which is actually made. Equally, some account is required of our ability to rehearse a number of verbal items silently. This is characterized in the model by the passage of information through the Response Buffer and back to the Logogen System. It is assumed that material in this buffer is organized serially and that it is coded in some form of articulatory code. Although there is no prior evidence which would motivate a precise assignment of the level of this code, there are reasons in terms of the internal consistency of the model which lead to the suggestion that the code is phonemic rather than phonetic.

Thus, at this point in the model, " pit " would be represented as /pit/ and " spit " as /spit/ rather than the *phonetic* forms [pʰit] and [sp⁼it] respectively, where [pʰ] and [p⁼] represent the aspirated and unaspirated forms of /p/. In the interests of economy of coding the specification of such allophonic variants would be left to later stages as would such aspects of speech as elisions between words and suprasegmental features such as intonation and stress assignment. Since it is required that the Response Buffer is implicated in silent rehearsal, it is necessary for the Logogen System to be able to receive information from it. The simplest assumption is that such inputs are specified in the same form as the outputs from the Logogens, i. e., a phonemic specification. The simplest possible system would be one in which input and output were identical and tied to one another.

Speech signals are analyzed by the Acoustic Analysis System (AAS) which indicates to the Logogen System the presence, in the signal, of any features which have been recognized. These features may be at any level of complexity from " the presence of fricative noise " to identification of specific syllables. One alternative pathway is required which links the Acoustic Analysis System to the Response Buffer without involving the Logogen System. This is required to account for our ability to read or mimic nonsense syllables for which there could be no representation in the Logogen System. The understanding of a word is mediated by a message sent from the Logogen System to the Cognitive System where the message is decoded and interpreted semantically. The basic form of the model has great simplifying power in describing phenomena of word recognition, the interaction of context with stimulus information (Morton, 1969*a*), short-term memory (Crowder and Morton, 1969 ; Morton and Holloway, 1970 ; Morton, 1970 ; Morton, Crowder, and Prussin, 1971), card sorting (Morton, 1969*b*), shadowing and delayed auditory feed-back. In addition, it yields quantitative predictions about the interaction between word frequency and threshold which are upheld compared with predictions arising from theories such as analysis-by-synthesis (Morton, 1968*b* ; see also Broadbent, 1967). It is of interest, then, to examine the way in which the existing model can handle the date concerning *A*'s behavior in the course of his phonological development.

THE DATA AND THE MODEL

Within the model there are two ways for the child to understand a spoken word. The first of these involves the normal path from AAS to the " acoustic " input of the Logogen System. The second involves the path from AAS to the Response Buffer and thence via the rehearsal loop to the phonological input of the Logogen System. We may suppose that the inputs from AAS to a Logogen will be determined by the adult form of the appropriate word. Thus /kwi:n/ will be recognized directly as " queen. " The child's own form of the word, /ki:m/, when analyzed by AAS, would not be accepted by any logogen. It could only be understood by the child mimicking the sound ; that is, by the analysis of the sound being fed round the alternative pathway, into the Response Buffer and then feeding the appropriate articulatory code back into the Logogen System. Since we have already assumed that a logogen is sensitive to a string which it produces as an output, this articulatory feedback will stimulate the logogen associated with " queen, " and thus enable the stimulus to be understood. Since these two recognition procedures are so widely different, it is reasonable that the child should react to them in different ways.

If a child is asked to recognize a word which is a homonym in his speech, we should expect that the normal recognition procedures would operate, leading to the adult form being understood. Since the question of what the sound means has been answered, the relative difficulty of prompting the alternative recognition system is comprehensible. For adults, a similar difficulty arises in trying to identify the six alternative interpretations of sentences such as " He filled the tank with petrol. " To give a somewhat more complex example, a person asked to say the string of words " whale, oil, bead, hammed " will be unaware of the alternative interpretation of the string of phonemes which most of the people listening to him would instantly apprehend.

We can equate the successive stages of development of the child's phonology with changes in the specification of the articulatory output from sets of logogens. This would be accompanied by changes in the specification of the acceptable inputs from the rehearsal loop. It follows then, in the context of this model, that as soon as the rules have changed there is no way in which the child could possibly understand one of the words produced as a result of his earlier rules. We would predict that exceptions to this statement would have occurred if adults, or other children, had regularly used A's incorrect form in speaking to him. In such cases we would expect a separate logogen to be set up with an input from AAS which accepted the variant which would then have the status of a synonym. While this was not the case with A, a wealth of informal consultants have reported that " baby talk " forms continue to be understood long after the children have ceased to use them.

Having dealt briefly with the way in which the model accounts for the primary observations, we will consider in more detail some of many outstanding problems.

THE SEPARATION OF ARTICULATION FROM ACOUSTIC ANALYSIS

We have seen that homonyms are created as a result of systematic misarticulation, and that in these cases A invariably recognizes the form that is correct by the adult model. It will be realized that analysis-by-synthesis theories of speech recognition, whereby words are recognized by reference

to internally generated signals, cannot handle such observations without severe modification. In particular there would have to be two kinds of generation, one which corresponded to the adult model for use in recognition, and a different one for speech production. Within the Logogen Model there is no direct connection from the articulation system to the speech-analysis system, the former being regarded as developing in a very different way from the latter. The need for this kind of separation can be illustrated by a couple of further examples from A's speech. At one stage he regularly pronounced *puddle* as *puggle*, this being the only environment in which this substitution was made. This is quite a common error for English children. The only way in which A could be persuaded to pronounce /p ʌ dl/ was by asking him to say *puzzle* when he regularly substituted /d/ for /z/. Thus we have a particular form, /p ʌ dl/, which means one thing when it is heard and another when it is spoken. This particular example would lead to an interesting paradox if we assumed that we monitor what we say by listening to and monitoring the acoustic side-tone. Thus, if A, intending to say *puzzle* and saying /p ʌ dl/, listened to his own acoustic output and checked it against his intentions, he ought to think that he had said *puddle* and correct himself. The corrected version would, of course, again be /p ʌ dl/... ! The only escapes from this loop are either to suppose that his auditory-recognition system operates in a totally different way when he is listening to his own speech (which might be plausible for vowel sounds but is scarcely so for consonants) [1] or by supposing that one's own speech is not monitored for its verbal content (cf. Morton, 1968*a*).

We can, in fact, give a more technical account of the child's difficulties with homonyms by taking the " sip " dialogue reported above as an example. There will be two relevant logogens, one defined semantically as *sip* and the other as *ship*. As the input specification of " ship " includes the /S/ which occurs in the adult form, the spoken word " sip " will not affect it, but will only trigger the *sip* logogen leading to appropriate semantic description being found and thence to " understanding. " When A is asked to find an alternative meaning it may be supposed that he uses the alternative pathway from AAS and the rehearsal loop, thus presenting an articulatory code /sip/ to the Logogen System. Such a code would be accepted equally by the *sip* and the *ship* logogens. However, it is well established that once a particular word has been used (recognized or generated), it has, for some time afterwards, a lower visual-duration threshold (e. g., Neisser, 1954). In the Logogen Model this fact is expressed by saying that once a Logogen has been active it requires less information to become active again. Thus, when /sip/ is fed round the rehearsal loop, the *sip* logogen will fire again before the *ship* logogen has time to respond. The *ship* logogen will thus be blocked. In the dialogue it needed the occurrence of a semantic cue, " it goes in the water, " to overcome this situation.

---

[1] Gives a greater dialectical variation among vowels than among consonants.

A further set of data which illustrates the separation of the recognition and generation systems concerns the very first contrast which A used to signal the voiced-voiceless distinction. This occurred in word-final position where all plosives were voiceless. Now, one of the perceptual cues to voicing in this position is that the vowel is longer when the final plosive is voiced. This is not distinctive, but is an allophonic variation entirely conditioned by the context. However, A used vowel length as his first method of signaling the voicing feature, thus contrasting bat vs. bad as [bat] vs. [ba:t] and *cart* vs. *card* as [g̊a:t] vs. [g̊a::t]. Smith tested the perceptual consequences of this by deliberately producing forms which violated the normal constraints on vowel length. This had no effect on A's understanding of the words, /ba:t/ being recognized as *bat* and /băd] as *bad* in spite of the cue of vowel length signaling the opposite interpretation in A's own speech.

Beresford (personal communication) has observed an equivalent case with /p/ and /b/ in word-initial position where there is constrast both in voicing and in force of articulation, /p/ being voiceless and with strong articulation while /b/ is voiced and with weak articulation. His patient, a 10-year-old boy distinguished between the sounds solely by force of articulation, both sounds being unvoiced. Thus *pin* and *bin* were produced as [pin] and [b̥in] respectively, where [b̥] represents the weak, voiceless bilabial plosive. Beresford tested the perceptual implications of these productions by presenting his patient with the form [b̊in] where [b̊] is voiced and with strong articulation. If the perceptual distinction followed the difference used in production then that sound should be understood as *pin*, following the force of articulation. In fact it was understood as *bin*. Thus the perceptual system must use the voicing characteristics of the sound.

HOW IS THE ACOUSTIC ANALYSIS SYSTEM SET UP ?

In the previous section we were suggesting that while the child is developing his productive phonology, the acoustic-analysis system is totally unaware of the activities of the articulators. Ultimately we will probably accept that this is too strong a hypothesis, but it will serve to focus on the problems. The question then arises of how the child ever learns the acoustic patterns of speech. We reject recourse to " innateness " hypotheses as nonproductive. It is, of course, possible that the infant's acoustic-analysis system is genetically endowed with structures which represent an analogue of the human vocal tract which automatically assigns phonological values to speech inputs. We prefer, however, to assume that what is innate is a set of generalized learning strategies which the infant brings to bear on any sensory input in order to seek out significant generalizations. We see as one of our future tasks that of specifying the nature of these strategies and tracing their progress. Beliefs that babies imitate the sounds in their environment seem

to beg the questions they aim to solve since, with the exception of features such as intonation patterns, imitation of speech requires knowledge of the vocal tract which is producing the model.

To understand the problem further, it is necessary to consider a little more carefully the tasks which a young child faces. For our present purposes we can distinguish between three activities : the formation of concepts concerning the world of objects and actions ; the recognition of particular sound patterns which relate to these concepts ; the production of vocal units which represent the concepts. It seems to us unnecessary to assume that all these activities are coordinated from the beginning. There is, for example, no reason to reject the possibility that concepts resembling the adult notions of " mother, " " food, " " chair " or " dog " should be relatively fully formed long before the possibility has occurred to the child that such concepts should be the referents of specific speech inputs. Equally, we have no reason to assume that the child should equate the acoustic consequences of his first babbling with the acoustic stimulation produced by the adults in his environment. At the moment we have neither the data nor the theoretical framework necessary to permit any precise statement of the possible interactions between the various functions which will later be involved in language. Let us then, for the moment, simplify the problem and assume that the effect of babbling is to set up in a part of the brain concerned with the analysis of acoustic inputs specific feature-recognition routines of the type required for speech recognition. These would include the picking out of specific peaks of energy in the frequency distribution, a differentiation of voiced and nonvoiced sounds (on the basis of relative time of onset of high and low frequency energy, vowel length, etc.), differentiation of aspiration, the acknowledgment that the presence of one sound can alter the manifestation of another sound—such that the /a/ sound in /bababa/ differs from that in /dadada/— and the fact that one of the invariants concerned with one class of sounds (the plosives) requires extrapolation of the frequency of a changing energy peak. We can assume, then, that by age of, say, nine months, the child is equipped to begin a systematic phonemic analysis of any language in his environment ; that he can produce a description of a speech input in terms of the features present ; that he can, if necessary, produce one of the class of sounds which contains any one feature via a path equivalent to the alternative path in Figure 1.

This suggestion as to the role of babbling does not seem to be too outrageous. The search for invariants in the babbled sounds can be interpreted as a specific instance of a general principle which babies must apply to all stimuli. Without such a principle we would be at a loss to account for the way in which the child acquires a body image or learns about the visual world. Motor control in babbling need not be differentiated from any other kind of self-exploring movements. For the search for invariants in the sound to be possible, it would be necessary for motor information to be fed to the auditory system ; but since we suppose that such activity is not under the control of that part of the system which will later be responsible for speech, our postulated separation between the articulation system and AAS does not create any anomaly.

THE SPECIFICATION OF PHONOLOGICAL FORMS

When a child is ready to begin its language learning, we assume that it is capable of distinguishing virtually all of the acoustic features which occur in the adult speech. Some learning process, which is not our immediate concern, then links a particular acoustic pattern to a concept. This is, in effect, the first stage in setting up a logogen. The form of the output, however, still has to be specified. This becomes a problem if we accept that speech is a rule-governed activity and that the child is set up, in effect, to search for phonemes. The main task for the child, then, becomes one of specifying which of the many features which are recognized in the adult form of a word are in fact relevant (i. e., " distinctive ") in what will become his native language. We suggest that this specification, the phonological output from a logogen, is initially incomplete and that any unspecified features are filled in by the application of general rules which the child has formulated concerning the phonetic realization of words. In terms of the present model, such rules would be applied in or beyond the Response Buffer.

The usefulness of this formulation is best illustrated with reference to a phenomenon in A's speech which Smith terms " restructuring. " This occurs in the few cases where the forms which A uses cannot be predicted from the general rules. These exceptions are rare, but where they occur

TABLE 2. A developmental sequence illustrating " restructuring. "

Adult form	*A*'s forms			
	Stage 1 Age - 2:5	Stage 2 2:6	Stage 3 2:8	Stage 4 2:11
one	wʌn	wʌn	wʌn	wʌn
sip	wip	tip	tip	tsip
same	weim	teim	teim	tse m
some	wʌm	wʌm	fʌm	fʌm
fish	wit	wit	fit	fit
foot	wʊt	wʊt	fʊt	fʊt
feather	—	tedə	tedə	tsedə
top	tôp	tôp	tôp	tôp
Rules	s → w	s → t	s → t	s → ts
	f → w	f → w	f → f	f → f

they seem to indicate that A has incorrectly characterized the phonology of the word. Two examples of this are embedded in Table 2, which covers four stages in A's development. At the first stage in the sequence an initial /s/, under the influence of consonant harmony, is rendered as /w/ where the final consonant in a CVC syllable is bilabial. At this stage an initial /f/ also becomes /w/ in all contexts. At the next stage the initial /s/ has become a /t/ for *sip*, *same* and all other similarly constructed words except for *some* and its compounds *someone*, *somebody*, *somewhere* and *something*. In all these cases the /s/ remains /w/. At the next stage of interest *some* has become /fʌm/ and its compounds follow suit. At the same stage an initial /f/ starts to be produced correctly. This seems to indicate that, in some sense, the initial phoneme of the word *some* is being characterized as /f/ which has not yet appeared in his speech. This error was maintained at the next stage when, for all other words, /s/ had become /ts/. In fact *some* was not correctly produced until A was 3.5, almost fine months after all other examples of an initial /s/ were corrected.

At the second stage in Table 2, the one exception to the rule which converted /f/ to /w/ was with *feather* which A produced as /tedə/. At this stage we might conclude that this represented the way in which he had characterized the word in his phonology. The later evidence, however, indicates that such is not the case, for at the fourth stage the word has become /tsedə/. If we look for a possible reason for this change we can see that, whereas an initial /t/ is being correctly reproduced, an initial /s/, which had been changed to /t/ previously has now become /ts/ regularly. The neatest *linguistic* description then is that A first characterized *feather* as /seðə/ in spite of the fact that the initial /s/ does not appear.

It would, of course, be inconsistent with the current form of the model to suggest that at Stage 2, the output of the *feather* logogen was /sedə/ and that of the *some* logogen was /fʌm/. If such were the case, then since we have suggested that the phonological input to a logogen from the Response Buffer has the same characteristics as its output specification, A should not be able to understand his own version of the two words /tedə/ and wʌm/. The evidence is, however, that he did understand these forms. (It is not known, however, whether or not he would have understood /sedə/ and /fʌm/ which would have been more conclusive evidence.) To avoid this difficulty, we suggest instead that the appropriate psychological description should involve a three-level tree. This could be characterized as in Figure 2 where the intermediate levels are written as phonemes and the successive stages of development are described as changes at particular points in the tree. Such a characterization would imply that the child has a complete but incorrect specification of the error words. We then have to consider how such erroneous specifications are corrected.

An alternative formulation in shown in Figure 3. In this we illustrate the possibility that the child has made an incomplete specification of the error words rather than an incorrect one. Any features which have been left unmarked would then be added at later levels according to certain prin-

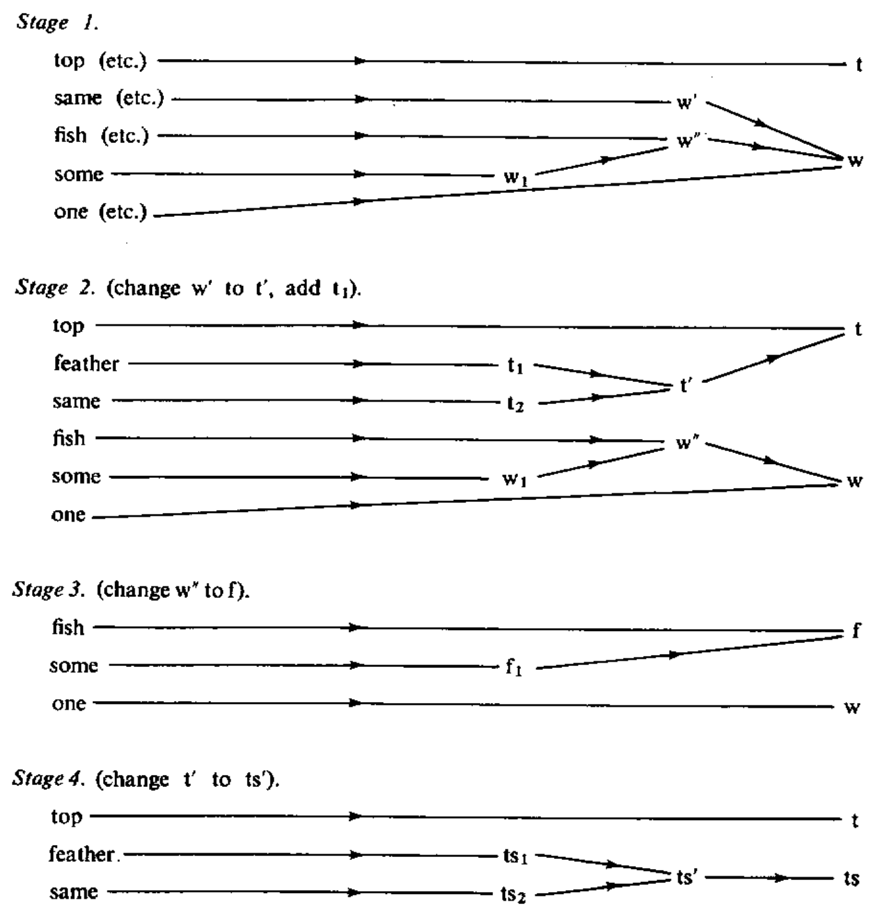
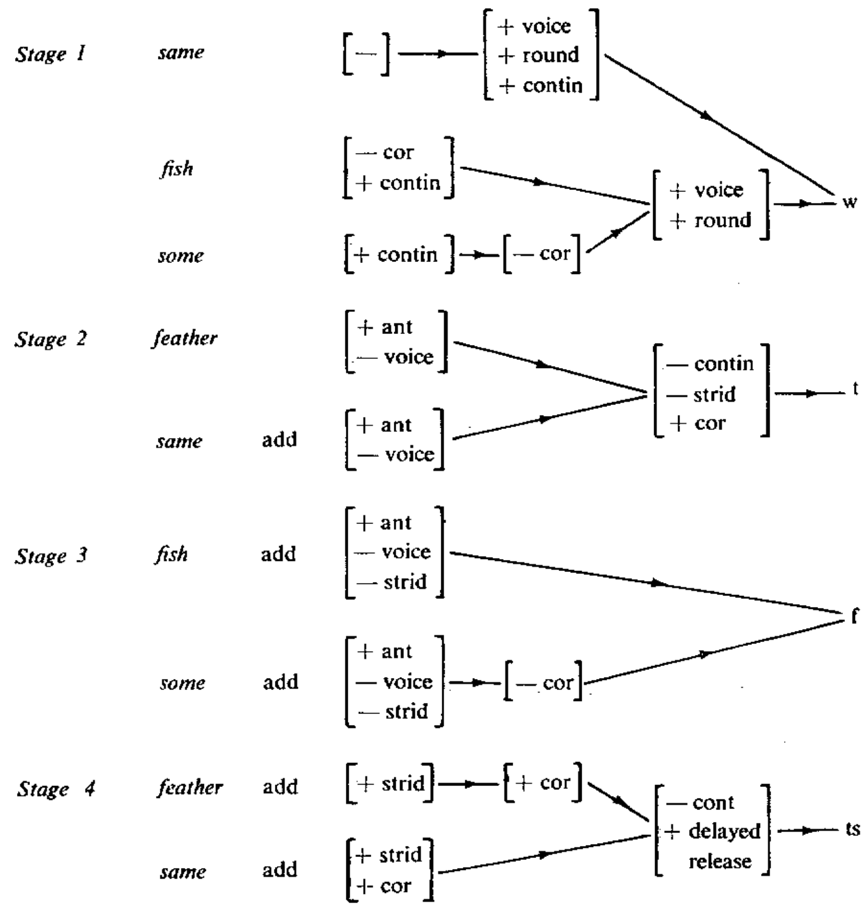


FIG. 2. A possible representation of the stages in Table 2.

Notes :

1. It is not intended to suggest, for example, that there are four kinds of *w* at stage 1. The various labels simply identify nodes in the tree.
2. The rules have the implication that if a node changes its 'name' then the connection to the right-hand side changes and all nodes connected to the left also change.



At later stages *feather* adds [+ cont] — *s* ; then [— cor] ; *some* adds [+ cor] ; *same* adds [+ cont].

FIG. 3. An alternative representation of the stages in Table 2.

Note : for *A*,/w/ is classed as [— voc, + cons] ; these features and [— nasal] are assumed throughout ; some other features are ignored. The sets of features on the left are those specified at the logogen level. Other features are added as a result of the phonological rules.

ciples (such as consonant harmony). As the child develops then, the specification of the phonological forms becomes more complete. Both of the representations have their drawbacks. The one in Figure 2, while economical in symbols, requires that we give some account of the mechanisms whereby the modes change. The second representations gets over the primary acquisition problem—all we need say is that the child continues to learn which features are relevant. However, we would have great difficulty, at this instant, in accounting for the peculiar phonological rules whereby additional features are specified.

It might be noted that, if the basic model correctly represents the processes involved in short-term memory, the correctness of the alternative formulations outlined above could be tested by examining the kinds of confusions made by a child when trying to recall a list of words. It is, alas, too late to perform the necessary experiments with A.

THE CORRECTION OF INCORRECT FORMS

At some stage it will be necessary to specify the mechanisms whereby the phonology changes. This we will not attempt to do. It is clear, however, that simple accounts of phonological development in terms of progressively more accurate imitation of the adult model is at least insufficient because children sometimes develop forms which do not occur in the adult speech. One example of this occured with A in conjunction with the acquisition of a systematic voicing contrast. At the first of the stages we will consider in this respect the child, A, pronounced all initial plosives as voiced and all final plosives as unvoiced. In addition /s/ was deleted in consonant clusters. Thus *spit*, *pit* and *bit* were all pronounced /bit/, *slug* was pronounced /lʌg/, *Smith* as /mit/ and *sweet* as /wi:t/. At the next stage he acquired the voiced-voiceless distinction and *pit* became /pit/. Simultaneously *spit* became /pit/ as well. Since /p/ and /b/ do not contrast in the context /s/ this need not have been expected. Smith, accordingly, suggests that the child has acquired the rule

$$\begin{bmatrix}+\text{cons}\\-\text{son}\end{bmatrix} \rightarrow [-\text{voice}]/\text{s}\_$$

Justification for this formulation occurred slightly later when the rule apparently became generalized to :

$$[+\text{cons}] \rightarrow [-\text{voice}]/\text{s}\_$$

The child then pronounced *Smith* as /m̥it/, with the voiceless nasal, and *sweet* as /fi:t/, /f/ being the nearest voiceless consonant to /w/. In addition, the cluster *sl* became reduced to /l̥/ so that the child contrasted *slug-lug* and *slip-lip*, not by the initial /s/, but by the absence or presence of voicing on the /l̥/.

Needless to say, neither /m̥/ nor /l̥/ had occurred in the speech to which A was exposed. Accounts of phonological development based on " successive approximation to the adult model " seem implausible compared with Smith's rule-governed explanation. Jacobson also reported some cases of phonemes being developed from outside the adult model. One example was that Russian children " who do not yet possess an *r*-sound often signal the presence of the /r/ by lengthening the vowel sound in the same syllable and temporarily create in this way a quantitative opposition which is otherwise completely unknown to Russian " (Jacobson, 1968, p. 14).

CODA

The foregoing is clearly only a small step towards a complete account of the acquisition of phonology. Apart from the simplification which we have indulged in, we need to specify the psychological reality of the linguistic rules which map the adult forms onto those of the child. While we have only just begun to explore the psychological significance of A's speech, we hope that others may be encouraged to perform equally detailed studies on other children.

REFERENCES

BROADBENT, D. E., Word frequency effect and response bias. *Psychological Review*, **74**, 1-15 (1967).

CROWDER, R. G., and MORTON, J., Precategorical acoustic storage (PAS). *Perception and Psychophysics*, **5**, 365-373 (1969).

JACOBSON, R., *Child Language, Aphasia and Phonological Universals* Mouton : The Hague (1968).

LIBERMAN, A. M., COOPER, F. S., HARRIS, K. S., MACNEILAGE, P. F., and STUDDERT-KENNEDY, M., Some observations on a model for speech perception. In W. Wathen-Dunn (ed.), *Models for the Perception of Speech and Visual Form.* Cambridge, Mass. : M. I. T. Press (1967).

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., and STUDDERT-KENNEDY, M., Perception of the speech code. *Psychological Review*, **74**, 431-461 (1967).

LIEBERMAN, P., Primate vocalisation and human linguistic ability. *Journal of the Acoustical Society of America*, **44**, 1574-1584 (1968).

MORTON, J., A preliminary functional model for language behaviour. *International Audiology*, **3**, 216-225 (1964) [Reprinted in R. C. Oldfield and J. C. Marshall (eds.), *Language*. London : Penguin Books (1968).]

MORTON, J., Grammar and computation in language behavior. In J. C. Catford (ed.), *Studies in Language and Language Behavior*, C. R. L. L. B. Progress Report No. VI, University of Michigan (1968*a*).

MORTON, J., A retest of the response-bias explanation of the word-frequency effect. *British Journal of Mathematical and Statistical Psychology*, **21**, 21-33 (1968*b*).

MORTON, J., The interaction of information in word recognition. *Psychological Review*, **76**, 165-178 (1969*a*).

MORTON, J., Categories of interference : verbal mediation and conflict in card sorting. *British Journal of Psychology*, **60**, 329-346 (1969*b*).

MORTON, J., A functional model for memory. In D. A. Norman (ed.), *Models for Human Memory*. Academic Press : New York (1970).
MORTON, J., CROWDER, R. G., and PRUSSIN, H. A. Experiments with the stimulus suffix effect. *Journal of Experimental Psychology Monograph*, **91** (1), 169-190 (1971).
MORTON, J., and HOLLOWAY, C. M., Absence of cross-modal « suffix effects » in short-term memory. *Quarterly Journal of Experimental Psychology*, **22**, 167-176 (1970).
NEISSER, U., An experimental distinction between perceptual process and verbal response. *Journal of Experimental Psychology*, **47**, 399-402 (1954).
SMITH, N. V., *The Acquisition of Phonology : A Case Study*. Cambridge : Cambridge University Press (1973).
STEVENS, K. N., On the relations between speech movements and speech perception. *Zeitschr. f. Phon.*, **21**, 102-106 (1968).

QUELQUES IDÉES SUR L'ACQUISITION DE LA PHONOLOGIE

J. MORTON et N. SMITH

RÉSUMÉ

La correspondance entre signaux acoustiques du langage et leur représentation phonémique est complexe. Les données de départ du modèle présenté ici sont dans l'étude que Smith a faite de la phonologie de son fils par des observations systématiques entre 2 ; 2 et 4 ; 0 ans. Les erreurs que cet enfant commet ne peuvent être attribuées à des problèmes perceptifs. L'évolution de ces erreurs est importante, car on peut montrer qu'elles impliquent des règles et leur application ordonnée. Ces règles se modifient au cours du développement pour arriver à une nouvelle prononciation qui est proche de celle de l'adulte. Il est important de noter que l'enfant ne comprend pas ses propres productions quand elles sont reproduites plus tard par un magnétophone, et que si elles sont homonymes de celles d'un adulte, alors c'est cette dernière forme qui est identifiée. On propose ici un modèle qui permet de rendre compte de la séparation entre articulation et analyse acoustique. Le concept central de ce modèle est celui d'un « logogène » qui a pour fonction de faire des mots, utilisables comme réponses. Chaque mot (ou morphème) est représenté par un logogène différent ; chacun de ceux-ci recevrait des autres éléments du modèle, une indication de la présence d'une information pertinente pour la production ou la reconnaissance du mot : quand l'information est suffisante, le logogène envoie un code au système cognitif ou au « buffer » de réponse. Dans la production, l'entrée du système logogène est le système cognitif (entrée sémantique) ; dans la reconnaissance l'entrée vient du système d'analyse acoustique (SAA). Le *buffer* de réponse a plusieurs fonctions : il doit séparer la réponse effectivement faite, la réponse disponible, la réponse répétée silencieusement ; ceci implique une boucle avec le système logogène (en codage phonémique). Le SAA indique au système logogène la présence dans le signal des traits reconnus ; ce système est aussi directement relié au *buffer* de réponse pour rendre compte de la possibilité de production ou d'imitation de syllabes sans signification qui ne peuvent avoir de représentation dans le système logogène. L'application de ce modèle aux données de Smith permet d'expliquer les différentes observations. Ce modèle est alors étendu à une description des débuts de l'acquisition du système phonologique par les bébés, acquisition dont l'imitation ne pourrait rendre compte.

Imprimé en France. — Imp. JOUVE, 17, rue du Louvre, 75001 Paris