8

CHILDREN'S LANGUAGE DEVELOPMENT AND ARTICULATORY BREAKDOWN[1]

Reprinted from:
D.L. Horton and J.J. Jenkins (Eds.)
The Perception of Language
Columbus, O., Chas. E. Merrill, 1971

8

CHILDREN'S LANGUAGE DEVELOPMENT AND ARTICULATORY BREAKDOWN[1]

KATHERINE S. HARRIS
Haskins Laboratories

Since speech production and perception appear to be almost exclusively human capacities, we are somewhat limited in our ability to examine their several substrata experimentally. Consequently, it has been common practice to try to discover the function of various parts of the system by examining pathological cases. These studies can be divided into two rough classes which we might call disability-oriented and normal function-oriented. Some workers are interested in diagnosis, prognosis and treatment problems; others are interested in what the study of the abnormal can tell us about normal language function. For some reason, although

[1]The work I have discussed here has been generously supported by the National Institute of Dental Research. The experimental work has been done in collaboration with Donald Shankweiler and Dorothy Huntington, and has been much more completely reported in the papers cited.

there have been studies of general language function which fall into both classes, studies of pathological phoneme production are generally directed towards treatment. We might be able to gain considerable insight into the function of various parts of the nervous system in controlling articulation, if we used clinical material on damage at different levels and sites in the sensorimotor system. This would be a large undertaking; in this paper we will only discuss some preliminary results on a special population, in particular, disordered articulation in patients with cortical damage, and, for comparison, in a population of patients with profound sensori-neural hearing loss.

Let us begin by positing a simple hypothesis about the nature of the phonemic disintegration which occurs when there is injury to some part of the nervous system responsible for phoneme production. Jakobson (1962) suggested that there is a hierarchy of difficulty in the sounds of speech; they appear in a regular order in the developing speech of the child, and that trauma causes the sounds to disappear in an order opposite to the order in which they first appeared.

There are a number of difficulties with this hypothesis, not the least of which is that we know little about children's language development. Furthermore, the original statements of the hypothesis are rather unspecific about what stage of language development is being discussed.

The usual textbook accounts (see, for example, Miller, 1951) of children's language development discuss it as divided into several stages: an early "cry" stage, followed by a "babble" stage, in which the variety of speech sounds supposedly increases rapidly until the child runs through a repertoire in which "all" speech sounds supposedly occur. At about one year, this stage is followed by the painful acquisition of words with a more limited repertoire, until, finally, the child at age three, four or five "correctly" produces words containing all sounds.

The simple textbook account is extremely difficult to verify. It is easy enough to make some acoustic measure of the developing speech of the child, but it is not clear what correspondence, if any, this form of transcription has to a phoneme transcription. Even for adult speech, there is no simple three-way-correspondence between the speaker's generation of a phoneme, some brief acoustic event and the linguist's transcription of a phoneme. The lack of correspondence can be demonstrated by trying to synthesize speech through abutting acoustic segments. This experiment has been shown by Harris (1953) to produce an unintelligible output. More sophisticated attempts to synthesize speech by rule have not been entirely satisfactory either. Even though a trained linguist can listen to a speaker and transcribe a string of phonemes, it is not entirely clear how he does it. Many engineers assume that the problem has been solved; for example, J. R. Pierce (1963) has written a very humorous article on the incorporation of a high-quality synthetic talker into complex human engineering systems. The problem, as he realizes, is that we know that humans talk and people listen, but it is not clear what they listen to. We cannot turn to a more "objective" technique which will be the equivalent of transcription. To make matters worse, the observer of a child's speech lacks the two advantages that the transcriber of adult speech possesses. First, he is making inferences about the speech gestures made in a vocal tract which is not only much smaller than an adult's, but may differ significantly from it in characteristics such as the relative size of the pharyngeal cavity. Apes differ from adult humans in this way (Lieberman, 1968), and infants may well be more like apes than like adults (Lieberman, Harris, and Wolff, 1968).

These shape differences, of course, will affect the acoustic output for a given speech gesture and may, therefore, confuse the listener. A second problem is that the field linguist can persuade his older informants to repeat a phoneme string, and to produce a contrasting string. This procedure is difficult with older children, and impossible with babies. Various workers in the children's speech areas have attempted to solve the problems in different ways—in some cases by ignoring their existence. Some representative approaches are indicated below.

One possible indication of the fact that a child "knows" a sound is that specially-trained observers consistently transcribe it in listening to his speech. This approach can be illustrated by referring to the work of Irwin and his various co-workers. Irwin was able to train observers to do a reasonably reliable job of transcribing the sounds of speech (Irwin, 1945), in that two of his listeners would transcribe the same string as containing the same sounds. The technique can also be used before the child is producing "real" words. However, the phonemes extracted by this method do not necessarily represent the same articulatory maneuvers that an adult makes in producing the "same" sound. For example, let us consider the "vowels" transcribed by Irwin, at various points in the developmental sequence (Irwin, 1948). He shows that in infants up to 5-6 months, the vowels transcribed are overwhelmingly ɪ, ɛ, and æ. His generalization from this is that front vowels appear before back vowels. However, it is well known from general acoustic theory that the form and frequencies of the vowels will depend on the overall length of the vocal tract. (See, for example, Stevens and House, 1961). The listener may be transcribing as fronted vowels those which are in fact produced with the tongue in middle position, because his ear has not compensated adequately for the difference in tract size in infants. Obviously, this point can only be checked by instrumental means, coupled with some precise data on infant vocal tract size.

A second approach can be used only if the child will repetitively produce "real" words on an elicitation procedure of some sort, as in the work of Templin (1957) or Morley (1957). The elicitation may be by pictures, or by having the child repeat the word after a model. Since the observer "knows" the target word, he can transcribe the produced phoneme string as containing, or not containing, the required sequence. To return to vowels, the child would only be considered to have learned the extreme front vowel, / i /, if it occurred in those words for which it occurs in adults. This is a much more stringent criterion than Irwin's and will yield data which show phoneme appearances much later in development, and not necessarily in the same order.

A third approach is an intermediate approach between the previous two. The sounds of speech are considered to be bundles of features. A child is considered as having or not having, a feature contrast or series of contrasts. Albright and Albright (1958), for example, describe a one-year-old child as having a five-vowel system. This means that there are five vowels which are contrastively used in words, but these five vowels are not necessarily the same as those of adults, in that they are produced by the same shaping of the vocal tract.

It is clear that these three techniques, because of the changes in criterion, will yield different results when we ask at what age a child "knows" a given phoneme. Furthermore, techniques one and three do not yield data on errors, whereas a technique like that of Templin, in which a word is elicited that can be compared with an adult production, will allow a classification of errors. Let us assume, for the moment, that we will consider only elicitation data on speech development. Do the sounds of speech appear in a child's speech in the reverse order to their disappearance in adult pathology? The technique used in the two studies we know of, our own (Shankweiler and Harris, 1966) and Fry's (1959), both involve having the patient repeat a standard list of utterances, a technique closely analogous to the techniques of Templin and Morley.

A first problem is whether or not the elicitation procedure itself will yield repeatable, reliable data. The Morley and Templin studies used similar procedures and apparently similar child populations. In general, the children produced a standard series of words, and these words were analyzed for phonemic content. Phonemes could then be scored on the relative percent correct in the population as a whole. We would expect the relative percents correct for the series of phonemes to be highly positively correlated from the two studies. In fact, the correlation, for consonants only, for percent correct by phoneme is only $r = +.43$, a rather low value. Of course, there are a number of obvious differences between the two studies. The words used, and the detailed elicitation procedures were different. The Morley children were English, the Templin children, American. However, if we believe that the data reflect a basic biological capacity, we would expect it to be stable from group to group. Fortunately, the Templin data allow us an opportunity to see if two groups are closely comparable under her procedure. She administered her test to a group of 3.5-year-olds and a group of 4-year-olds. The correlation between scores is $r = +.95$. It is possible, then, to obtain stable data; the low correlation between Templin's and Morley's results must be due to factors other than fundamental instability. However, we obviously need some extremely painstaking work in this area to find out what are the relevant variables.

We come now to a comparison of the child and dysarthric data. Of course, the suspicions we have of the stability of data from elicitation studies affect any conclusions we might draw about the relationship between the child data and any other. However, if we forge ahead, we can correlate the Templin child data with data from an experiment in which the subjects were brain-damaged and, consequently, had dysarthric speech. The nature of the subject population and the experimental procedure have been previously reported in greater detail (Shankweiler and Harris, 1966). The correlation between the percent correct performance for various phonemes in the two groups is $r = +.70$ if we compare it with the Templin study, or $r = +.78$ if we compare it with the Morley study. These are extremely substantial values, especially in view of the possible instability of the basic data.

Apparently, then, there is some substantial basis for the notion that some consonant sounds are more difficult than others and consequently more likely to deteriorate when the production system is damaged. However, a more careful examination of the production problems of the child and adult dysarthric population suggests that things are not this simple. Templin does not give detailed data on error classification, but does remark (p. 55):

> Whenever a sound element was not produced correctly, the inaccuracy was classified into the following categories: Omissions, defective sounds, and substitutions . . . For the sample as a whole, substitutions were approximately 10 times as frequent as omissions, and about 4.5 times as frequent as the use of defective sounds.

In our own listening to the dysarthric speakers, we forced a transcription — that is, we did not allow the "defective" category. However, one interpretation of the difference between "defective" and "substituted" sounds is one of reliability of transcription. Templin's listeners had apparently no trouble classifying a substantial portion of the errors as another standard sound. In our experiment, we tended to transcribe errors as long, frequently non-English strings, but these transcriptions were not reliable from one observer to another. This might mean that the category, "defective sounds," is some substantial order of magnitude larger for dysarthrics than children. This impression is confirmed by listening to the production tapes. The same sounds seem to give trouble for the two groups, but the nature of the error is different — the dysarthrics just do not sound like children who can't quite articulate yet.

The category of voicing errors is worth special comment and has already been noted by Fry (1959). Voicing errors are rare among children, according to Morley. Menyuk (1968), in a recently published study, concludes that the voicing distinction develops quite early. Voicing errors are common in adult dysarthrics, as shown in both our study and Fry's. Again, this point argues against a simple disintegration hypothesis.

Thus far, we have been considering only consonant sounds. However, it is our feeling that we would not come to very different conclusions if we had included vowels. Children make fewer vowel errors, and so do dysarthrics. Our conclusion would have to be that simple error data give modest support to Jakobson's hypothesis, but that the nature of the errors suggests that the lack of coordination in the dysarthric is not simply primitive, but different in kind from that of a child. It is interesting to compare this data with that from a group of deaf speakers (Huntington and Harris, 1968). If we assume that certain sounds are inherently "harder" than others, then we might suppose that harder sounds would be

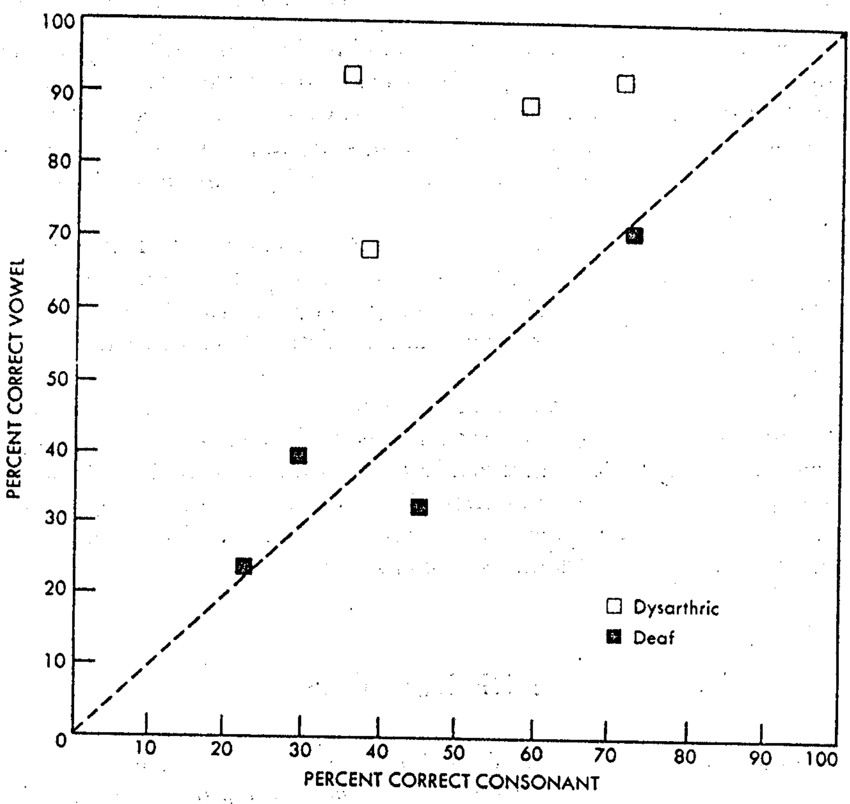

FIGURE 1. Percentage of vowel and consonant errors in the deaf and dysarthric groups.

harder to teach and that, consequently, deaf speakers would have trouble with the same sounds that are hard for children. Again, we compared deaf speakers with Templin's and with Morley's children. The correlations for order of difficulty of consonants are +.58 and +.20, respectively, which are not as high as the child to dysarthric correlation, but still suggest that some of the same problems are involved. However, again we find that the nature of the errors is different. Voicing control is difficult for deaf speakers, as has been known since the classic study of Hudgins and Numbers (1942).

However, the most notable difference between deaf speakers and the other groups is in the prevalence of vowel errors. Figure 1 shows the relative percentage of vowel and consonant errors in the deaf and dysarthric groups. It is easy to see that deaf speakers have markedly more trouble with vowels. "Hard" and "easy" sounds are different for the two groups. Again, this seems to argue against a simple one-factor primitivisation hypothesis for articulatory breakdown.

I would like to conclude by summarizing the points I have tried to make in this somewhat discursive presentation. First, we can gain some insight into the complex problem of articulation control and development by carefully examining the cases of articulatory insufficiency.

The highly inadequate data presently available suggest that different pathologies lead to rather specific patterns of deficit, and that these patterns cannot really be considered to be like those of a child who does not yet articulate correctly. Furthermore, our account of the development of articulation in normal children is itself in need of considerable further study.

REFERENCES

Albright, R. W., & Albright, J. B. Application of descriptive linguistics to child language. *Journal of Speech and Hearing Research*, 1958, 3, 257-61.

Fry, D. B. Phonemic substitutions in an aphasic patient. *Language and Speech*, 1959, 2, 52-61.

Harris, C. M. A study of the building blocks in speech. *Journal of the Acoustical Society of America*, 1953, 25, 962-69.

Hudgins, C. V., & Numbers, F. C. An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, 1942, 25, 289-392.

Huntington, D. A., Harris, K. S., & Sholes, G. N. An electromyographic study of consonant articulation in hearing-impaired and normal speakers. *Journal of Speech and Hearing Research*, 1968, 11, 147-58.

Irwin, O. C. Reliability of infant speech sound data. *Journal of Speech Disorders*, 1945, 10, 227-35.

Irwin, O. C. Infant speech: development of vowel sounds. *Journal of Speech and Hearing Disorders*, 1948, 13, 31-34.

Jakobson, R. Kindersprache. In *Selected writings*. Netherlands: Mouton, 1962, 328-401.

Lieberman, P. Primate vocalization and human linguistic ability. *Journal of the Acoustical Society of America*, 1968, 44, 1574-84.

Lieberman, P., Harris, K. S., & Wolff, P. Newborn infant cry in relation to nonhuman primate vocalizations. *Journal of the Acoustical Society of America*, 1968, 44, 365 (A). Manuscript in preparation.

Menyuk, P. The role of distinctive features in children's acquisition of phonology. *Journal of Speech and Hearing Research*, 1968, 11, 138-46.

Miller, G. A. *Language and communication*, New York: McGraw-Hill, 1951.

Morley, M. E. *The development and disorders of speech in childhood.* Edinburgh and London, 1957.

Pierce, John R. The paper dragon: a tale of the time. *Physics Today*, 1963, 16, 45-50.

Shankweiler, D., & Harris, K. S. An experimental approach to the problem of articulation in aphasia. *Cortex*, 1966, 2, 277-92.

Stevens, K. N., & House, A. S. An acoustical theory of vowel production. *Journal of Speech and Hearing Research*, 1961, 4, 303-20.

Templin. M. C. *Certain language skills of children*. Minneapolis: University of Minnesota Press, 1957.