An Experimental Approach to the Problem of Articulation in Aphasia*

Donald Shankweiler† and Katherine S. Harris
Haskins Laboratories, New York City

The purpose of this paper is to outline an experimental approach to an old problem: the nature of the articulatory disturbance which is often associated with so-called "motor" aphasia. It has long been known that in some cases of disturbance of expressive language, impairment of speech production occurs at the most molecular level; that is to say, the patient's difficulties are not in finding words, but in producing the gestures words are made of. Patients who suffer from the condition Alajouanine has called "phonetic disintegration" articulate slowly, hesitantly, with extrinsic facial movements and many substitutions of one phoneme for another. This condition, which is sometimes called "cortical dysarthria" (Bay, 1962), must be distinguished from dysarthric disorders arising from damage at lower levels of the motor system. In the latter, articulatory movements are impaired but the neural mechanisms responsible for organization of the gestures which convey the phonological units of speech are not involved.

Phonetic disintegration can occur in combination with aphasic disturbances or in remarkably pure form. Apraxia of the oral-facial region is usually present, but it is not always com-

* Some of the material on which this paper is based was presented at the annual meeting of the Academy of Aphasia, Niagara Falls, Canada, October 1965.

† Also at the Institute of Physical Medicine and Rehabilitation, New York University.
parable in severity with the articulatory disturbance. Therefore, the possibility has to be considered that the anatomical arrangements for speech do not overlap completely with those which serve other cranial movements (Geschwind, 1965; DeRenzi, Pieczuro, and Vignolo, 1966).

The reawakening of interest in the disorders of articulation associated with aphasia owes much to the work of Alajouanine and his colleagues. Alajouanine, Ombredane, and Durand (1939) argue forcefully for the necessity of obtaining an adequate phonetic description of speech in cases with disturbed articulation, viewing this as an indispensable requirement for understanding the physiological bases of phonetic disintegration. Important suggestions are made about the ways speech production is linked to the various stages of recovery, ranging from the earliest stage in which only vowels can be produced to later ones in which virtually all segments of speech can be identified, though they are produced unreliably and in the wrong articulatory contexts. Unfortunately, the results of their phonetic studies are not presented in such a way that the reader can make his own evaluation of their findings. The work also suffers from failure to provide the necessary controls.

Fry (1958) has provided an illustration of a method for systematic analysis of the utterances of an aphasic patient with articulatory disorder. The problem was to develop an adequate test to enable an investigator to determine the relative difficulty of all the sounds of speech in their various articulatory contexts. Not only must the sample include all the words required for a complete survey, but it must include enough words of each type to provide adequate sampling. The problem was met by having the patient repeat a list of words chosen to meet these requirements. Phonetic analysis of the patient's tape-recorded utterances was carried out and a complete tabulation of all responses was made.
This permitted a classification of all substituted phonemes according to place and manner of articulation. Fry's paper demonstrates the relevance to aphasia studies of a powerful method for discovering the dimensions of a phonological disturbance.

Mention must be made of another approach to the study of articulatory disorders. Lehiste (1965) has carried out meticulous spectrographic analyses of the utterances of a diverse group of patients suffering from a variety of neuromotor disorders, many of them due to damage to brain stem and cerebellar regions. We cannot compare her findings directly with our own, because we did not use the technique of spectrographic analysis and because we confined our study to patients with presumably unilateral lesions of the left cerebral hemisphere.

The preliminary findings presented here are a part of a larger ongoing comparative study of articulatory function in cases of damage at different levels and sites in the sensorimotor system. In the present study we have tried to gain an accurate picture of the dimensions of phonemic error which occur in the syndrome of phonetic disintegration. Information regarding the consistency and major directions of phonemic substitutions can then serve to guide more direct approaches to the defective speech gestures by means of electromyography, x-ray cineradiography and other techniques. Our program has these two interrelated aspects: first to obtain an adequate assessment of articulatory function in phonetic terms and secondly to study directly the gestures of speech in order to relate the phonetic description of the disordered speech output to the parameters of muscle movement. The present study is limited to the first aspect.

Methods and Results

Subjects

We have made phonetic studies of the speech of five patients of the Institute of Physical Medicine and Rehabilitation of New
York University. Each had suffered a stroke six months to seven years prior to the examination. The stroke in each case was followed by severe expressive aphasia with preservation of comprehension. At the time of testing, aphasic symptoms were greatly diminished, leaving a major residual deficit in articulation.

The patients' ages at the time of testing ranged from 39-61 years; three were under 50. None presented difficulties of chewing or swallowing, nor showed any tendency to drool. Some weakness of the right upper extremity was present in all patients, and in two the leg was also involved. Slight weakness of the right side of the face could be detected. All were able to perform movements of the lips, face, and tongue on command, but with obvious apraxic impairment. No cerebellar signs were present and there was no evidence of lower motor neuron paralysis. In two cases the episode could be traced to occlusion of the left internal carotid artery. In a third case, angiography showed occlusion of the middle cerebral artery, and in two cases the origin of the lesion is not definitely known.

Although very different degrees of recovery of speech are to be found in the group, the sequence of events during recovery was similar for all. Writing was spontaneously adopted as the way around obstacles to oral communication. With the return of speech, the difficulties proved to be chiefly of an articulatory nature. All patients had reached a definite plateau in recovery at the time of testing. The residual level of articulatory impairment ranged from gross difficulty in producing all classes of speech sounds, with poor intelligibility, to mild impairment shown mainly in production of consonant clusters, with good preservation of intelligibility. None showed evidence of generalized intellectual deterioration.

Much interaction over a long period of time had convinced us on clinical grounds that the patients were intact in compre-
hension of speech. A control, however, is needed in order to be reasonably certain that the errors in articulation occur on the basis of impaired motor organization, rather than on the basis of perceptual impairment. A "rhyme" test was used to determine whether the acoustic cues which signal phonemic differences could be utilized normally by these patients.

**Assessment of Speech Perception**

Each patient listened to a taped recording of three 75 word lists of real word monosyllables. He was to encircle each word he heard in a set of five possible answers for that word printed on an answer sheet. For example, if the word spoken was *pat*, the answer sheet might offer the alternatives *pat, mat, slat, bat, cat*. In one-third of the list, the alternatives differed, as in the example given above, in initial sound; one third of the answer sets varied in terminal sound, and one third in vowel. Thus, we tested the patients for the ability to perceive initial, terminal and vowel sounds in words. We made up the words for the vowel part of the test; the initial and terminal section lists were taken from a test described by House et al (1965). As a further control, the same test was given to six speech therapists. The results for four patients (one left therapy before we could administer the test) and six therapists are shown in Fig. 1.

Two patients performed almost as well as the therapists; indeed, if we had used unselected members of a normal population as controls, we might have expected comparable results. Two performed less well; showing, however, the same pattern: greatest accuracy on initial sounds, more errors on terminal sounds, and poorest performance on vowels. It will be shown below that this is the reverse of the order of accuracy in producing speech. Moreover, the patient (JG) who showed the greatest impairment in articulation (see Fig. 3) was virtually as accurate as the therapist controls in perception. Consequently, it is unlikely
that difficulties in selecting the correct sound from the multiple-choice sets could account for the errors in production. The errors on the rhyme test being largely confined to vowels -- which have the most irregular spelling of the three classes of sounds -- are more consistent with reading difficulty than with impairment of auditory perception.

Assessment of Articulation

Since the principal need was to study speech production at the phonological level, the investigation was limited to analysis of the productions of single words. A test consisting of 200 real word monosyllables was assembled. The distribution of phonemes in the word list was such as could be expected to bring into relief the difficulties in articulating the speech sounds of English. The list contains most singleton consonants, a sample of the most frequently occurring consonant clusters, and those vowels which in the regional dialect are not ordinarily characterized by glides. One respect in which this test differs from articulation tests designed for other purposes is in providing equal frequency of occurrence of all phonemes in each position in the word. Each of 25 consonants and clusters occurred eight times in initial position, similarly, each of a set of 25 consonants and clusters occurred eight times in terminal position, and each of eight vowel nuclei occurred approximately 25 times in the list. Even distribution of phonemes is highly desirable if one wishes to draw inferences about the relative difficulty of different speech sounds. The words were recorded on magnetic tape by a trained speaker and presented to the patient through earphones at a comfortable listening level. The patient's task was to repeat each word once. The responses were tape-recorded so as to be available for detailed analysis.

The results of articulation testing were treated in the following manner. The patients' recorded utterances were
transcribed by a phonetically-trained listener. The transcription can best be described as phonemic. The intent was not to make the transcription reflect every phonetic nuance of speech production, for the practical reason that this would involve more categories than could be dealt with without obscuring the main features of the condition. A very broad transcription admittedly ignores features of speech which may have real importance for a physiological understanding of phonetic disintegration, but it has the virtue of allowing the gross features to stand in bold relief. Therefore, we have chosen to ignore narrowly phonetic features in this exploratory phase of our work.

The transcribed utterances for each patient were tabulated as confusion matrices; two such matrices appear as Fig. 2. The matrix shows, with respect to each phoneme, the frequency with which it was correctly produced or replaced by another phoneme. Each row in the matrix refers to the phonemes occurring in the list, and each column refers to the sounds actually spoken by the patient. Since each initial consonant occurred eight times in the list, the figure 8 in the diagonal indicates that that phoneme was repeated correctly on every occasion it occurred. From the diagonal we can quickly discover the relative difficulty of the phonemes, and from the other cells in the matrix, we can discover the degree to which the phonemic substitutions fall into patterns.

The matrix of initial consonant substitutions for two patients who represent the extremes of the group, is shown in Fig. 2. Clinical judgment had placed HL as the least impaired in the group, and the data bear this out. Her errors in articulation were largely confined to a few consonant clusters and certain fricatives (/θ,v,z,dʒ/). Clusters involving /l/ were produced in a stereotyped manner by splitting the cluster and inserting
a vowel in between, as /pəlɪz/ and /kəlin/ thus producing disyllables. The patient lacked sufficient control to produce the cluster as a unitary gesture and had been taught this substitution by her therapist as a way of producing words in which these sounds occurred intelligibly if defectively. The lower matrix in the figure shows initial consonant confusions in the most severely impaired patient in the group. There is a striking tendency to use voicing inappropriately. This is best illustrated by tracing the occurrences of three voiced consonants /b,d,g/. It will be seen that one or another of these sounds was substituted repeatedly for almost every other sound in the inventory.

Matrices of this type were made for each patient, one for initial consonants and clusters (as the examples shown in Fig. 2), one for vowel nuclei and one for terminal consonants and clusters. These matrices, in turn, could be summarized by considering the overall percent correct for each matrix, shown in Fig. 3.

It is apparent from Fig. 3 that the errors in articulation are not evenly distributed throughout the word. This figure shows the tabulation of the total number of sounds correctly produced in each position in the word for every patient. Not surprisingly, consonant clusters are more difficult to produce adequately than singleton consonants. In every case consonant clusters in the initial portion of the word are the least accurately produced of all sounds. One patient (JG) failed to produce any of these sounds intelligibly. In four of the five cases, we can make the generalization that initial sounds are more difficult to produce than terminal sounds. The vocalic portion of the word is produced with greater accuracy than the nonvocalic portions. This is particularly striking in the case of JP.
Assessment of the production of vowels introduces some special problems owing to the fact that differences in vowel quality are the main sources of phonetic variability between speakers of the same language. In view of this, the listener's assessments of the patients' vowels are more tentative than identification of other portions of the speech signal. In many instances there was uncertainty as to whether a certain feature should be attributed to pathology or to the fact that there may be several "correct" vowels for any given word. This ambiguity could be resolved by comparing each patient's productions against those of a close relative who is healthy, preferably a sibling of about the same age.

Patient-sibling comparisons were made in two cases (HL and JP) by retabulating vowel production using the vowels of their siblings as targets. When this is done, HL's percentage accuracy rises from 97% to 99% and JP's figure changes from 81% to 88%. Although these changes are small, they are a further indication that the production of vowel sounds is very little affected by the disorder. Although the fifth patient, JG, has poor vowel production, we might expect some upward revision if a sibling control were available.

Considering the five cases as a group, there is a great deal of variability in the pattern of sound substitutions which occur, but some sounds were consistently misarticulated. Fig. 4 shows the sounds which the group as a whole found difficult to produce. It includes all phonemes in the inventory for which the wrong sound was produced on at least three of its eight occurrences by at least four of the five patients. The number to the right of each phoneme is the mean per cent error in producing it. It is noteworthy that the lists contain fewer sounds in final position than in initial, an additional indication that these patients have greater difficulty in producing
<table>
<thead>
<tr>
<th>Phoneme</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>62</td>
</tr>
<tr>
<td>f</td>
<td>58</td>
</tr>
<tr>
<td>s</td>
<td>55</td>
</tr>
<tr>
<td>lk</td>
<td>47</td>
</tr>
<tr>
<td>ps</td>
<td>65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>75</td>
</tr>
<tr>
<td>v</td>
<td>63</td>
</tr>
<tr>
<td>z</td>
<td>65</td>
</tr>
<tr>
<td>tf</td>
<td>55</td>
</tr>
<tr>
<td>d3</td>
<td>55</td>
</tr>
<tr>
<td>pl</td>
<td>82</td>
</tr>
<tr>
<td>kl</td>
<td>73</td>
</tr>
<tr>
<td>sm</td>
<td>65</td>
</tr>
</tbody>
</table>
initial sounds. The fricatives and affricates, together with certain consonant clusters, are the classes of speech sounds which are most consistently misarticulated.

The fricatives and affricates are also the most difficult sounds to perceive correctly. Confusions are common, particularly between /f/, /θ/, and /v/ (Miller and Nicely, 1955). Therefore we cannot ignore the possibility that some of these errors in production may reflect perceptual confusion, compounded with the likelihood that the transcription also is somewhat unreliable in dealing with sounds of this class. Both possibilities can be checked to some extent by retabulation. Consonant sounds may be classified in terms of the place in the oral cavity at which maximal constriction occurs in forming them (e.g., /b/ is labial; /g/ is velar), and the manner in which they are produced (e.g., /p/ is plosive; /n/ is nasal). The errors were classified according to whether the substituted sound differed from the target in place of articulation, manner of articulation, both place and manner, or whether the substituted sounds were unrelated to the target or omitted altogether. When the productions of stop consonants and fricatives were retabulated allowing manner errors, the error rate for fricatives increased whereas that for stop consonants decreased. Thus the fricatives remain the most impaired class of sounds when the tabulation allows as correct those substitutions which are likely to occur on a perceptual basis.

The results of the tabulations for all classes of speech sounds show a similar incidence of errors of place and manner with manner errors slightly predominating. The largest class, which accounted for one-third of all errors, was the category of unrelated substitutions and omissions.
Discussion

A feature of phonetic disintegration which is shared with many other disorders of articulation is the special difficulty in producing fricative and affricate sounds. Luchsinger and Arnold (1965) have observed that one fricative /s/ is the most likely to be impaired in any type of pathological speech whatever the native language of the speaker. The fricative sounds require the use of more muscles and closer control of the amount and timing of movement than, for example, the stop consonants. They are among the last to be added to the child's repertoire of speech sounds (Whetnall and Fry, 1964).

The finding that the fricatives, affricates and some consonant clusters are the only sounds which are consistently misarticulated by patients of our group points to the conclusion that no particular structure or region can be implicated to the exclusion of other parts of the articulatory apparatus. Sharply localized defects of specific muscles or muscle groups were not to be expected in a disorder of this kind. The possibility of dissociated defects of the lips and tongue has been raised (Denny-Brown, 1958), but we have found no evidence of differential involvement of these structures in any patient of our group. The preservation of vowel articulation, which was also observed by Nathan (1947), argues against any selective impairment of the tongue. So does the absence of general impairment of the lingual consonants.

The hypothesis that the most rapid gestures of speech suffer most in phonetic disintegration was suggested by the fact that vowel production, which requires a relatively slow rate of movement, was least disturbed by the disorder. The implication that production of the stop consonants (requiring the fastest movement) would be most impaired is contrary, however, to what we found.
To what extent do we generally find substitutions of easy sounds for difficult ones? This question has proved to be a stumbling block in considerations of the nature of phonetic disintegration. One reason is the ambiguity of what is meant by phonological difficulty. A relatively small proportion of the errors represent phonetic simplifications typical of young children (such as substitution of stop consonants for fricatives). A second class of errors must also be regarded as simplifications although they do not occur in children's speech (such as the breaking up of consonant clusters by insertion of a vowel between the normally linked pair of consonants). A third class of substitutions cannot be considered simplifications in any sense. On the contrary, the patients often emitted particularly difficult strings of consonant clusters which do not occur in English words and were nearly impossible to imitate.

Our findings, therefore give little support to the idea (Jakobson, 1941; Alajouanine and Lhermitte, 1960) that speech in phonetic disintegration mirrors its development in the child. In this we are in agreement with Critchley (1952) and Fry (1958). It is true that our patients find difficult some of the same sounds the young child finds difficult (Morley, 1957), but there is little reason to assume that the difficulty has the same basis.

Phonetic disintegration does not resemble other disorders of articulation in a number of important ways. Children with developmental articulatory defect and adults with defects of the articulatory structures are quite predictable in the substitutions they make, whereas variability is one of the striking features of the speech of our group. The errors do not group systematically according to place of articulation. For example, there is no significant tendency for front consonants such as /p/ to be better produced than middle or back consonants such as
/t/ or /k/. A number of errors occurred in the manner of articulation. Chief among these were errors of voicing, particularly the replacement of an unvoiced sound by a voiced one. This type of substitution rarely occurs in children's speech (Fry, 1958).

Many errors can be considered neither as poor approximations to the correct phoneme nor as clear-cut substitutions of related phonemes. The apparent unrelatedness of many of the substituted sounds to their targets, together with the marked lack of consistency of the substitutions tells us much about the condition. To the extent that these features characterize a given patient's performance, the disorder must stem from disorganization of the process by which phonological units are encoded for production. It is almost inconceivable that residual spasticity or weakness could give rise to errors of this kind.

We have not attempted to describe all pathological features of speech in cases of phonetic disintegration. Certainly the patients have difficulties in producing connected speech which are not revealed by the method of analyzing monosyllables spoken in isolation. Syllabification, in particular, has been indicated as a problem in need of study. Lehiste (1965) and Tikofsky and his co-workers (1965) have shown that the intelligibility of utterances produced by their group of dysarthric patients varied with syllable content.

Even within the present restricted framework, there are a number of questions which cannot be answered by studies of the acoustic speech output alone, whether by the trained listener or by spectrographic analysis. Neither can give direct information about the motor gestures of speech. New developments in the application of electromyography to speech research (Harris et al, 1964) make it possible to obtain a description of phonetic disintegration directly in terms of the disordered
movements. An application of EMG to the study of a case of developmental dysarthria was illustrated by Rootes and MacNeilage (in press), and we are hopeful that this technique, used in conjunction with the method outlined here, can clarify a number of remaining questions. We would like to know, for example, whether the vowel gestures are as little affected by the disorder as the identification data would suggest. To what degree, if any, is bilateral asymmetry present in the action of the tongue, lips and palate? What is the contribution of residual dystonia to the errors of articulation?

It is of basic interest to learn whether or not all the normal components of each distinctive gestural unit of productive speech are present in the utterances of these patients, both when the target sound is acceptably produced and when another is substituted for it. If the gestures are intact in form, what errors of timing occur? In the normal workings of speech all participant processes are so thoroughly fused in their integrated action that the relative contribution of each cannot be understood. We believe that careful examination of these questions in persons with localized interruption of a motor or sensory link in the speech chain can contribute new understanding to the processes of speech production.

Summary

A phonetic analysis was made of speech production in five patients with major residual deficits in articulation following recovery from expressive aphasia. The findings demonstrate major disturbance of speech production at the most molecular level. Maximal difficulty in articulation occurred at the beginning portion of words. Consonant sounds were much more often misarticulated than vowel sounds. Fricative and affricate consonants and certain linked groups of consonants were the most often affected of all sounds of speech. Phonetic simplifications
typical of young children were observed less frequently than other errors which are not found in children's speech nor in the speech of adults with defects of the articulatory structures.

Systematic observation of speech in phonetic disintegration can contribute new understanding of the fundamental processes of speech production.

Acknowledgments

We are indebted to Martha L. Taylor, Chief, Department of Speech and Hearing Therapy, Institute of Physical Medicine and Rehabilitation, New York University, for providing the impetus for this study and for her considerable help and encouragement. Thanks are due to Ida Nathan for administering some of the tests and to Sheila Zelemeyer for preparation of the confusion matrices.

The work was supported in part by a Rehabilitation Research and Training Center Grant from the Vocational Rehabilitation Administration, U.S. Department of Health, Education and Welfare to the Institute of Physical Medicine and Rehabilitation, New York University, and in part by a General Research Support Grant from the National Institutes of Health to Haskins Laboratories.
References


Tikofsky, R.S. *Phonetic Characteristics of dysarthria*. (Office of Research Administration, Univ. Mich., 1965.)

Figure Legends

Fig. 1. Speech perception test: per cent correct identification of initial consonants, terminal consonants, and vowels; shown individually for each of four patients and averaged for six therapists.

Fig. 2. Confusion matrices for initial consonants and consonant clusters for two patients. Each line of the matrix shows the distribution of sounds the patient produced for the normal consonant indicated on the left margin. Cells forming the diagonal are "correct" productions. Symbols follow usage of the International Phonetic Association. Approximate pronunciations as follows: θ as in thin; ʃ as in shoe; θ as in then; ʧ as in chin; dʒ as in just.

Fig. 3. Speech production test: per cent correct for each indicated class of sounds for each patient.

Fig. 4. Speech sounds with highest error rates given as percentages of total opportunity for error.