Current Status of Electromyography in Speech Research*

Thomas Gay†
Haskins Laboratories, New York City

There are several reasons why electromyography is particularly suited to the study of speech. Unlike x-ray motion pictures, for example, which supply only indirect information about speech movement, EMG enables us to look directly at the muscle actions responsible for this movement. In this sense, we can study the dynamics of speech production, not only in descriptive terms of what muscles are contracting and when, but moreover, the relationships between the component gestures of speech and their linguistic counterparts. It follows naturally, too, that information of this type is equally relevant for the study of both normal and defective speech production.

This paper consists of a brief description of the current instrumental methods used for obtaining and processing electromyographic data and the particular applications and problems associated with their use for studying speech.

The functional unit of a muscle is known as a motor unit. This is a structure comprised of a single nerve fiber and, for the

† Also, Hunter College of the City University of New York
Fig. 1. Muscle action potentials recorded at weak effort (A and B) and strong effort (C)
muscles of speech at least, a hundred or more individual muscle fibers. When a neural impulse arrives at the motor end plates of these fibers, a brief contraction, in the order of only a few milliseconds, results. Accompanying this contraction is a small electrical potential which is dissipated into the surrounding tissue. By placing a pair of electrodes near the muscle, these potentials can be picked up and displayed on the face of an oscilloscope.

Figure 1 illustrates a typical recording of muscle action potentials. The top trace shows a succession of potentials generated by a single motor unit. As the force of contraction increases slightly (middle trace), a second motor unit is called into play and finally, a much stronger contraction (bottom trace) involves the action of many more motor units and accordingly, shows a greater overall amount of electrical activity.

A major difficulty in recording the muscle activity of speech is that these muscles are too often difficult to isolate. This problem, of course, can be somewhat obviated by the use of needle electrodes. Conventional needles however, present certain drawbacks. Their rigidity and relatively heavy weight can cause movement artifacts and sometimes dislodgment not to mention the considerable amount of subject discomfort experienced. Placement procedures are often complicated and finally, it must be assumed that the particular motor unit at the electrode accurately reflects the actions of the other motor units of the muscle. Some of these problems however, seem to have been overcome recently by Shipp (1968) and his colleagues who successfully studied the deep muscles of the neck using a wire electrode system. This system, originally developed by Basmajian and Spring in 1955 (and I believe independently by Hirano, in Japan), consists of a thin wire threaded through the cannula of a hypodermic needle. The exposed end of the wire is bent back over the needle to form a hook. The needle is used only to carry the wire electrode to the desired location
after which it is withdrawn. Upon withdrawal of the needle, the hook portion of the wire becomes anchored to the muscle. Removal of the wire is accomplished by a slight tug. Shipp reports that the wire is flexible enough to permit natural movement and that the procedure is relatively painless. One problem with this system however, is the possibility of the hook breaking off during removal. The chances of this happening though would depend on the type of wire used and the location of the muscle. In addition, the practical difficulties of placement associated with needle electrodes, are also encountered with the hooks.

These problems, inherent with needles and to some extent, the hook system, have led many investigators to consider the use of surface electrodes when interference from adjacent muscle activity is not a factor. Unlike needles, surface electrodes are easily placed and afford minimal discomfort. In addition, whereas needle electrodes pick up the activity of generally a single motor unit, surface electrodes sum over many motor units and thus, provide an indication of overall muscle activity, a somewhat more useful measure for studying the speech gesture. In order to obtain a convenient quantitative record of this activity however, the raw EMG signal is customarily transformed into a display of amplitude versus time. This is easily accomplished by rectifying and integrating the original signal.

Figure 2 shows the same signal in both its original and integrated forms. Generally speaking, the peak height of the integrated curve is an indication of the strength of the muscle contraction. This is only an approximation however, since the amplitude of the recorded signal varies with the distance between the electrodes and the active motor units. Further, since the integrated curve represents the vector sum of a number of uncorrelated muscle potentials and since productions of identical utterances vary from one token to the next, a number of curves must be

129
averaged before we can arrive at a reasonably accurate picture of the muscle activity at a given electrode position.

The basic data processing procedure then, is to collect myographic data for a number of tokens of a given utterance or utterances and to computer average the integrated signals for each electrode position. This general technique has proved quite satisfactory for a number of experiments performed by several research groups, notably those at the Department of Linguistics, U.C.L.A. and at Haskins Laboratories.

A block diagram of the EMG system presently used at Haskins is shown in Figure 3. The overall procedure involved three separate operations: first, the myographic data are recorded on one-inch magnetic tape; second, the data are played back on a pen writer where each utterance is labeled and edited; and finally, the tape is played to the computer for the actual averaging.

The system contains sixteen data channels of which eight are for the recording of the myographic signals. The surface electrodes are of the active suction type, connected to a manifold, which in turn is driven by a vacuum pump. The other inputs are for the acoustic signal (one each for air and throat microphones), channels for strain gauge and air pressure data, a banter channel for the operator's comments and finally, three channels for a calibration signal, a clock pulse and an octal coded pulse. All data inputs are recorded on a sixteen channel magnetic tape recorder. During the recording part of the experiment, a record is also made on an eight channel pen writer. This serves to monitor both the EMG and acoustic signals. There are also provisions for visually monitoring the gain levels and auditorally monitoring any of the recorded signals.

In order to identify each utterance for the computer, a code pulse, preceding each utterance, is also laid down on the tape. This pulse is in octal code and can be triggered either manually or automatically by the second channel of a cue tape. The first
KSH 25 March, 1965
Item 436 DAH
List XV, Item 4
Code 713

1 UL

2 CL

3 LL

4 BT

5 MT

6 TT

7 Code (oo per sec)

8 Voice Trace

Start Count-down Start Sampling Line-up Point Stop

[p] closure and friction

Termination of phonation

Duration of Token

Code 713 Item 436 [hə'paɪk]
channel of this tape has the test utterances recorded for presentation to the subject. Along with the EMG and acoustic data and the utterance identifying code pulses, a calibration signal is also recorded. This signal is fed to the computer during the processing run for purposes of obtaining a final readout in microvolts.

Figure 4 illustrates the editing operation. The first six traces show a re-run of the integrated EMG signal. The output from the throat microphone is plotted on track 8 and the code pulse, superimposed on a 50 Hz clock pulse, is written out on track 7. The decision is first made as to how the different tokens of an utterance are to be lined up in time with respect to the averaging process. This is then measured as the distance from the code pulse to the pre-selected line-up or time-zero point which in this illustration is the offset of the utterance. However, this point can be positioned anywhere along the utterance as can the "start" and "stop" averaging points.

During the data processing run, the tape is played continuously to the computer where all calculating and tabulating operations are done automatically. The print-out of the curve is in microvolts at sampling points as short as five milliseconds. Standard deviation estimates are also included in the program.

Figure 5 shows a typical set of curves derived from the mean voltages at each of the sampling points. The curves are referenced against various points along the duration of the utterance. The line-up point in this case is the onset of voicing.

At present, all editing is done by hand but plans for the near future include a more automated procedure. In fact, the hardware is now operational with only the program yet to be written. The new procedure will allow all data to be stored on a disk file after which each utterance will be retrieved and displayed in digital form on the face of an oscilloscope. Here, a potentiometer control adjusts the location of the line-up point. After the
averaging process, the data will again be displayed on the scope, this time in analog form, where the voltage curve can be either traced or photographed.

The system to date has proved quite satisfactory for its intended purpose and moreover, is capable of recording data from needle electrodes or, with a few adjustments, from other electrophysiological inputs.

I would like to add as a final note that certain technical problems still remain with EMG. Because of their overlapping locations, various muscles of the speech mechanism cannot yet be accurately mapped. Even with reliable locations, procedures for electrode implantation are not always straightforward; and finally, the processing of myographic data is often a laborious task. Complications notwithstanding though, I think it apparent that the technique of electromyography is a significant and prominent tool in physiological speech research.

References
