THURSDAY, 25 JULY. MORNING

JOINT SESSION WITH THE INTERNATIONAL SOCIETY OF EXPERIMENTAL PHONETICS

Chairman: Prof. R. H. STETSON

44. Dr R. CURRY (Newcastle): *The cathode-ray oscillograph in speech recording, with special reference to the study of the modern Northern English dialects.*

In beginning I do not propose to devote any time to a description of the cathode-ray oscillograph, since I have tried to explain the principle of the oscillograph during the course of my demonstration. The particular make of oscillograph I have used is a medium-voltage, gas-focused model having a photographic type of fluorescent screen. This oscillograph is manufactured by Messrs A. C. Cossor Ltd. of London. I should like to mention the following details concerning the method of operating the oscillograph. I have used electrostatic deflection with the deflector plates for two reasons. Firstly, the high impedance of the deflector plates enabled me to use valves of high impedance and high amplification factor in my amplifier; and secondly, the field of the deflector plates is screened by the surrounding field of the extended anode and so the oscillograph is less sensitive to external fields. Moreover, I have used a R.C.C. amplifier with a dynamic microphone, calibrated by the makers, and the response of which I have been able to calibrate subsequently with the aid of a heterodyne low-frequency oscillator and sound-source. In recording with moving film I used 16 mm. sound-recording film manufactured by Messrs Ilford Ltd. of London. I used a special camera in which the film in strips 1 metre in length was mounted on the periphery of a disc which was rotated at a constant speed of 5 feet per second past the lens. The exposure was controlled by the speaker, who before speaking pressed a switch to begin the exposure and terminate it after one complete revolution of the film disc. This method was evolved to enable us to economize in film, which was a considerable item of expenditure. As you will understand, an average recording for each speaker consisted of some thirty records, each requiring 3 feet of film and the whole costing 10*s.* 6*d.* The method has certain disadvantages, yet these are balanced by considerable advantages which I will describe later.

Before giving details of the actual technique of recording and the information gained from the oscillograph records, I should like to point out some of the most obvious superiorities of the cathode-ray oscillograph over other types of oscillographs. In the first place it is generally acknowledged to be the most accurate method of recording alternating phenomena at frequencies up to the finite speed of the electron beam. Recent investigation by Dr MCGREGOR MORRIS has shown that any distortion due to the residual gas in the tubes does not approach 1 per cent. at frequencies up to 25 kilocycles. With the more recent tubes developed for television purposes

this distortion is absent. These tubes are suitable for very fast film recording and may be operated at anode potentials of 6000 volts. Furthermore, the cathode-ray oscillograph can be tested for response to either frequency, amplitude or phase changes. I have no time to devote to these tests. Dr DAVIS has shown that distortion of amplitude and phase may be large in recording with Duddell or Einthoven oscillographs when the recorded frequency approaches even a quarter of the resonant frequency of the vibrator. From more general considerations the cathode-ray oscillograph is much more easy to handle —the only fragile part being the filament or cathode as compared with the delicate suspension of the vibrator. I have used the Einthoven oscillograph and I know how delicate the suspension can be. To the best of my knowledge it is impossible to damage the cathode-ray oscillograph with an overload from the microphone.

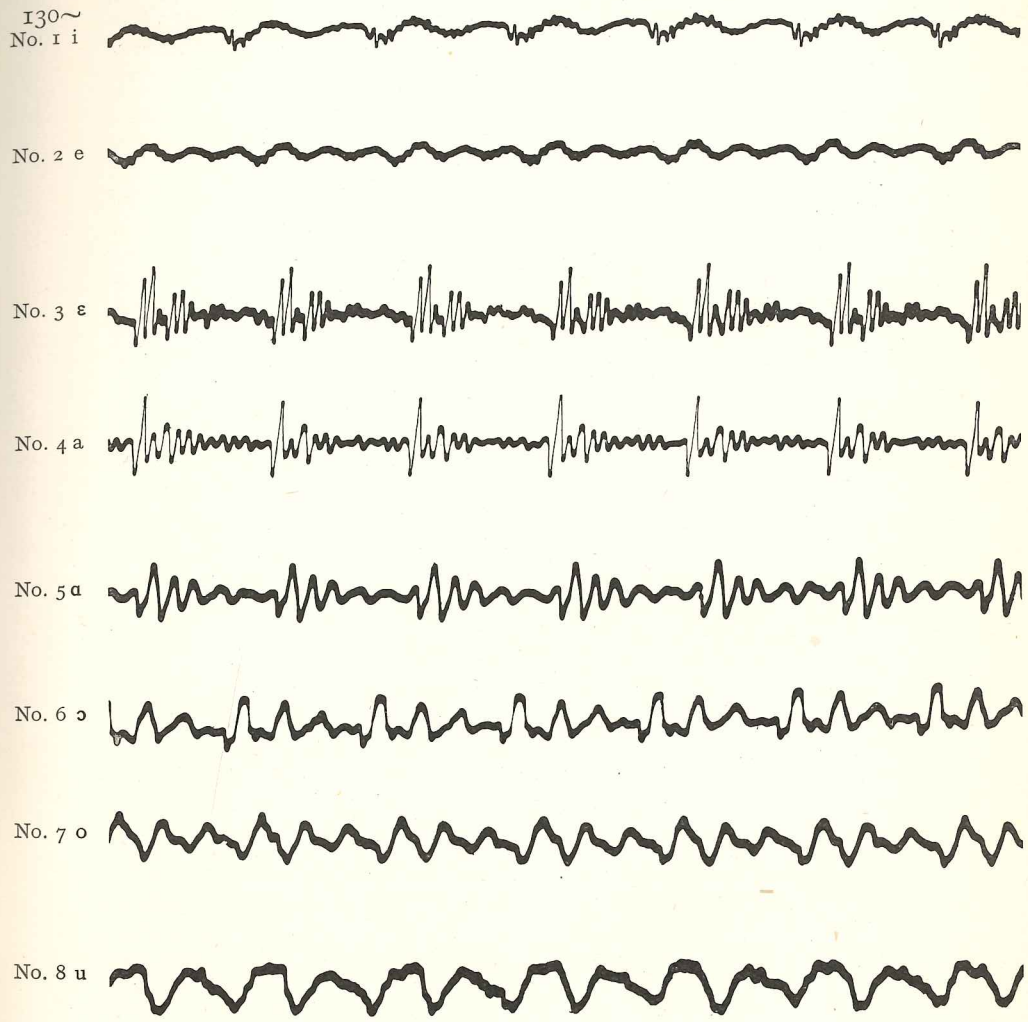
A second great advantage possessed by the cathode-ray oscillograph is the facility with which the wave-form of any periodic oscillation recorded on the oscillograph can be made visible as a pattern on the fluorescent screen. This is achieved by coupling the microphone circuit to the vertical deflector plates and at the same time coupling to the horizontal plates an electrical circuit of special design. By means of this circuit the microphone impulses give a vertical deflection which is spread out horizontally at a uniform and controllable speed. When the frequency of the horizontal sweep coincides with the fundamental period of the maintained vowel or voiced consonant, the wave-form of the vowel or consonant appears as a stationary pattern on the fluorescent screen. This is an obvious advantage over other types of oscillographs, which can produce similar results only with the aid of rotating mirrors which are usually difficult to operate. The electrical circuit used to produce the horizontal delineation of the wave-form is generally based upon a fundamental circuit consisting of a condenser and resistance connected in series across a source of high potential and coupled to the deflector plates of the oscillograph. Various refinements upon the original circuit are used—notably, variable-mu S.G. valves in place of the resistance and a Thyratron or gas-discharge triode across the deflector plates. The purpose of these additions is to ensure that the circuit operates to charge the condenser at a uniform rate and discharge it rapidly before the saturation point is reached. This circuit produces a saw-tooth wave-form which permits of a faithful delineation of the wave-form of the microphone impulses. The value of this method of visual delineation of the speech wave-forms is very considerable, and as will be appreciated it is possible to demonstrate all manner of phonetic problems to a group of students and then allow them to try and find their own individual characteristics of speech.

In describing now the technique of making the film records, I should like to point out that we have been handicapped by two limiting factors. Firstly, we were anxious to avoid any considerable wastage of film, and consequently determined to limit each record to a film strip 1 metre in length and recording at most a word of three

syllables. Secondly, we preferred to make a series of records of separate words rather than to record a phrase or sentence on a long length of film. These limitations were of course inter-related. The disadvantage was that the speaker was required to operate the recording mechanism himself at a time when he felt confident of being able to pronounce the word correctly. The advantages were, firstly, that we were able to make the speaker practise the word before he made the record, and secondly, that we had the opportunity of noting down in phonetic script the actual pronunciation recorded on the film strip. These are considerable advantages from the linguistic point of view, since it will be understood that in the case of a dialect which is considerably influenced by standard spoken English the speakers do not always give the true dialect pronunciation at the first attempt. Moreover, the simultaneous phonetic record provides an excellent test of one's power of making phonetic analyses by ear only. As I have mentioned already, we make on an average thirty records with each speaker and we have over 400 such records already made and under examination. In some records the words were prefixed to a phrase the whole of which was not recorded. The total group of words afforded examples of the developments in the dialect from the Middle English sound system. The words were very carefully chosen for this purpose, and it has proved very instructive to compare the film records for the same word pronounced by speakers from several different areas. It was necessary to bring our speakers to the laboratory for these records, since our apparatus is not of a portable nature. The relative speed and facility of the method can be shown by the fact that we average thirty records in 45 minutes. This involves the loading and unloading of the film drum for each record. The records are developed and fixed later. Each record is labelled, dated and examined in comparison with the phonetic record made during the recording. Prof. Daniel Jones of London paid us a visit to see the apparatus and made records of the eight cardinal vowels, the wave-forms for which are shown in the figure which accompanies this description. These records were made on an average voice pitch of 130 cycles per second. One fact that stands out most prominently from an examination of the film records is the inadequacy of ordinary phonetic notation to represent the information conveyed by the film records. The customary length-marks are insufficient to convey the possible differences in length between the vowels and the elements of diphthongs. In records of diphthongs the process of transition from one element of the diphthong to another was usually clearly visible.

Finally, I should like to give a few details of our method of analysis. It was humanly impossible to subject each record to analysis whether by Fourier or Vercelli system. Accordingly, we had to compromise by making records of a group of vowels from the Modern Northern English dialects, which we used as our type wave-forms. These have been analysed, though I am not prepared to discuss whether such analyses are of great value. The wave-forms have been illustrated in the *Archives néerlandaises de phonétique*, XI, 107–18.

Our subsequent records have been examined largely by comparison with these standards. Two important facts should be mentioned in this connexion. In the first place it is possible to compare visually only those records which have been recorded on approximately the same voice pitch. Secondly, it is important to bear in mind the type of wave-form under examination. Thus I have found three main types of vowel wave-forms, viz. the **ɑː** type, the **uː** type, and the **iː** type.



Wave-forms for Cardinal Vowels recorded by Prof. Daniel Jones

From the linguistic point of view the Modern Northern English dialects include a complex range of vowels and diphthongs, which are not generally used in Standard English. The range of diphthongs is very wide. Quite considerable differences in dialect will be found between the speech of inhabitants of villages separated by only a few miles, and the field of investigation is very extensive. It gives me great pleasure to state that we have always found our speakers particularly anxious to help and I should like to thank them for their courtesy.

45. Prof. G. Oscar Russell (Ohio): *Synchronized X-ray, oscillograph, sound and movie experiments, showing the fallacy of vowel triangle and open-closed theories.*

It is to be hoped that all of you were fortunate enough to see the very fine talking X-ray movie of Prof. Menzerath. You noticed how movements of the velum, larynx, and tongue could be readily followed. And that the tongue was not the only organ occupied in creating vowel quality differences. Furthermore, that the back throat was obviously fully as much, if not more, involved than the front mouth.

So again we have ample proof that the unfortunate physiological front mouth vowel triangle does not represent the facts. This opportunity should not be permitted to pass, therefore, without calling attention to the necessity of the linguistic scholar and teacher adopting a more reliable classification scheme.

Since it must still be said we know practically nothing as to the physiological cause of fine vowel distinctions, and certainly that the physiological act is far more complex than the mere front tongue arching represented by that vowel triangle, the folly of holding to, and fallacy of the latter, must be obvious. But what can be substituted? That is the inevitable question. And a proper one. Let us consider it a moment.

Sound change is dependent on what we hear. Not on what our tongues feel. And the *normal learning process is also guided by the auditory* sense rather than by what the tongue feels. Where we have to rely on the latter even when supplemented by the visual, as in teaching the deaf, the process is so slow, and the results so inaccurate, as to make that very apparent. Since what we hear is obviously primary, our classification schemes, transcription characters, and terminology, should also be acoustic. For it is just a wildly unscientific absurdity to listen to a speaker's strange sound, and then proceed to record it on paper in terms of a physiological character and scheme. Especially where the latter are fantastic, and unsupported by every X-ray and other scientific experiment we carry out; being based, as they are, solely on a physiology which was originally, and is still, purely imaginary.

It should of course be recognized by all that the ear will not hear many fine distinctions in sound, which a high speed scientific experiment would record. The same thing holds true of the eye. It cannot see the flight of a bullet; hence if you want to know just what

happens, a high-speed motion picture of from 3000 to 90,000 exposures per second is more reliable. And it will tell more than the most voluminous argument about, and description of, what one, or a group, imagines the eye can, or should, see. But in that realm we understand the manifestations shown on the experiment. Whereas when we look at the experimental analyses of the vowel we, as yet, understand but little of what we see. The phenomena are so complex, and are dependent on so many misunderstood processes, that we stand baffled before our experiment once we have made it.

Where this is not true there are of course many details we can establish with far more reliability than any number of listening recorders could do depending solely on their ears. Is the vowel nasal, for example? Or is it partially unvoiced? Does it start before the consonant ends? And so on *ad infinitum.* It is sheer folly, in such cases, to reject the aid careful scientific experiment could lend, and to depend solely on a series of letters to record what the ear hears. That would be like the chemist who in this day and age attempted to write a treatise on acids and their physiological reactions, by recording in chemical and other learned symbols just what his tongue tasted and other senses perceived. That day is long since gone. And real scholars and scientists would in this day and age laugh such a procedure to scorn. For the available technique makes far greater accuracy possible.

Generally speaking we can say that a careful scientific analysis of all consonant manifestations is as much called for as in the case of the chemical problem above. For we do know what we see on those experiments. Is it a stop, fricative, velar, partially voiced or nasal? etc. He who relies on his ear when he can so much more accurately establish the facts by simple scientific experiments is to-day as far behind the times as the above-mentioned antiquated physiological chemist.

We also regularly detect that a given e is more "closed" than some i. Unfortunately, though, that is exactly what the linguistic scholar does not want us to find. For he is thinking in terms of what he hears and describing in terms of a physiology he imagines, which in actuality is wide of the mark. Obviously, then, the wise thing for him, and you, and me, is to cease that scientifically absurd process and describe, as well as think, in terms of what we all three hear. Then when we classify, use likewise a scheme which is based on the acoustic, rather than the antiquated imaginary physiological manifestations which are actually non-existent.

The cardinal vowel device of Prof. Daniel Jones is to that end recommendable. The phonograph record makes it possible for anybody to compare the vowel being considered with the reference norms given thereon. We have made such phenomenal progress in phonograph recording during the last decade that such an acoustic "yard-stick" is now quite reliable. Of course it should be re-recorded at intervals so as to keep it up to date as recording techniques are improved (assuming of course that the old are always kept available for verification of pronunciation uniformity). And it is my per-