The Sounds of Speech

By IRVING B. CRANDALL

NOTE: As professor of vocal physiology, Alexander Graham Bell did pioneer research in "devising methods of exhibiting the vibrations of sounds optically." In 1873, he became familiar with the phonautograph, developed by Scott and Koenig in 1859, and with the manometric capsule, developed by Koenig in 1862. Greatly impressed by the success of these instruments "to reproduce to the eye those details of sound vibration that produce in our ears the sensation we term timbre, or quality of sound" Bell used an improved form of the phonautograph having a stylus of wood about a foot long. He obtained "large and very beautiful tracings of the vibrations of the air of vowel sounds" upon a smoked glass.

In describing his early attempts to improve the methods and apparatus for making speech waves visible and to interpret wave form, Bell wrote:

"I then sang the same vowels, in the same way, into the mouth-piece of the manometric capsule, and compared the tracings of the phonautograph with the flame-undulations visible in the mirror. The shapes of the vibrations obtained in the two ways were not exactly identical, and I came to the conclusion that the phonautograph would require considerable modification to be adapted to my purpose. The membrane was loaded by being attached to a long lever, and the bristle, too, at the end of the lever, seemed to have a definite rate of vibration of its own. These facts led me to imagine that the true form of vibration characteristic of the sounds of speech had been distorted in the phonautograph by the instrumentalities employed. I therefore made many experiments to improve the construction of the instrument. I constructed, at home, quite a number of different forms of phonautographs, using membranes of different diameters and thicknesses, and of different materials, and changing the shape of the attached lever and bristle."

Struck by the likeness of the phonautograph and the mechanism of the human ear, Bell conceived the idea of making an instrument modeled after the pattern of the ear, thinking it would probably produce more accurate tracings of speech vibrations. In 1874, he consulted a distinguished aurist, Dr. Clarence Blake of Boston, who suggested that instead of trying to make an instrument modeled after the human ear, the human ear itself be used. Dr. Blake prepared a specimen containing the membrane of tympanum with two bones attached, the malleus and incus. The other bone, the stapes, was removed and a stylus of wheat straw about one inch long was substituted. A sort of speaking tube was arranged to take the place of the outer ear. "When a person sang or spoke to this ear, I was delighted to observe the vibrations of all the parts and the style of hay vibrated with such amplitude as to enable me to obtain tracings of the vibrations on smoked glass."

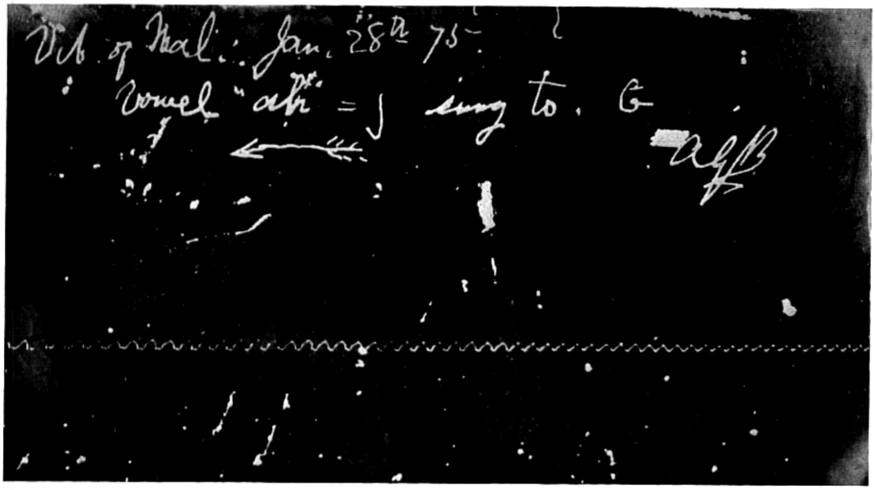
In the accompanying paper, Dr. I. B. Crandall describes modern methods whereby with the most refined apparatus, highly accurate speech wave forms have been produced. The analysis and interpretation of both vowel and consonant sounds made possible by these records, are the realization of an objective sought by Bell a half century ago.

This article is the result of an extended study of 160 graphical records of vowel and consonant sounds, of which a few are reproduced in the present publication. One hundred and four of these records are of vowel sounds and formed the basis of the "Dynamical Study of the Vowel Sounds," by I. B. Crandall and C. F. Sacia which was published in this Journal in April, 1924. The purpose of the present article is to describe all of the records in sufficient detail, including in one discussion the outstanding characteristics of vowel, semi-vowel and consonant sounds; it is hoped shortly to supplement this with a reproduction of a larger group of records from the complete collection.—*Editor.*

CONTENTS

INTRODUCTION

TO the layman speech is a matter of course, but to the student of science, or of language "the amazing phenomenon of articulate speech comes home . . . as a kind of commonplace miracle."[1] Hence we have inquiries into the nature of speech from many points of view, beginning with fundamentals based on physiology and acoustic science and leading to important applications in communication engineering, phonetics and vocal music.



Speech record made by Bell in 1875

The scientific study of speech sounds began with Helmholtz, who also made a fundamental study of hearing. Helmholtz had the advantage, in approaching these problems, of a knowledge of physiology as well as a mastery of theoretical physics. With this equipment and such simple laboratory apparatus as he created, he did his great work on speech and hearing of which we have the record (in English translation) under the title of "Sensations of Tone."[2] Today, with

---

[1] Greenough & Kittredge, "Words and Their Ways," N. Y., 1901.

[2] "The Sensations of Tone as a Physiological Basis for the Study of Music." Translated from the Fourth German Edition by A. J. Ellis: Fourth English Edition, London, 1912.

immeasurably superior physical apparatus, and with more specialized theoretical equipment, the individual investigator usually approaches one problem at a time, the problem and the method being selected according to the technique with which he is familiar. The work of D. C. Miller on sound and sound analysis[3] represents the beginning of modern physical research on speech sounds. In medical science some attention has been given to the mechanism of speech[4] and the psychologists are responsible for an enormous literature on voice control and the perception of speech and tones.[5] The work of Scripture[6] represents the beginning of a science of experimental phonetics, and in the closely related field of philology there is a rapidly growing interest in the physical characteristics of speech sounds.[7]

In this large field of investigation the physicist finds a real opportunity in providing means for the study and measurement of speech sounds, and a real responsibility in broadening the extent and improving the accuracy of such quantitative data as are obtained.

The results obtained from such physical investigations have practical as well as scientific value, and we observe that in a large laboratory concerned entirely with the development of electrical communication considerable effort has been devoted to research on speech and acoustic apparatus.[8] It has recently been felt that the wave

---

[3] "The Science of Musical Sounds," New York, 1916. This contains a bibliography of 90 special references, some 12 of which relate specifically to speech.

[4] "A Contribution to the Mechanism of Articulate Speech," by S. W. Carruthers. Edin. Med. Jour. VIII (New Series) (1900) pp. 236, 332, 426.

[5] "The Psychology of Sound," by Henry J. Watt (Cambridge, England, 1917), contains a bibliography of 159 references. The work of C. E. Seashore is noteworthy in this field.

[6] "Researches in Experimental Phonetics." Publication No. 44, Carnegie Institution, Washington, 1906.

[7] "The Physical Characteristics of Speech Sound," by Mark H. Liddell. Bulletin No. 16, Purdue University Engineering Experiment Station.

[8] See following papers, from the Research Laboratories of the American Telephone and Telegraph Co. and Western Electric Co., Inc.:

(a) H. D. Arnold and I. B. Crandall: The Thermophone as a Precision Source of Sound: Phys. Rev. 10, (1917), p. 22.
(b) E. C. Wente: Condenser Transmitter for Measurement of Sound Intensity: Phys. Rev. 10 (1917), p. 39.
(c) I. B. Crandall: The Air Damped Vibrating System: Phys. Rev. 11 (1918), p. 449.
(d) I. B. Crandall: The Composition of Speech: Phys. Rev. 10 (1917), p. 74.
(e) R. L. Wegel: Theory of Telephone Receivers: J. A. I. E. E. 40 (1921).
(f) E. C. Wente: Sensitivity and Precision of the Electrostatic Transmitter: Phys. Rev. 19 (1922), p. 498.
(g) I. B. Crandall and D. Mackenzie: Analysis of the Energy Distribution in Speech: Phys. Rev. 19 (1922), p. 221.
(h) H. Fletcher: The Nature of Speech and its Interpretation: J. Franklin Inst. 193 (1922), p. 729.
(i) J. Q. Stewart: An Electrical Analogue of the Vocal Organs: Nature, Sept. 2, 1922.

forms of the speech sounds required more precise determination, and indeed research in the art of telephony has emphasized this need. The graphical records of speech sounds, which form a supplement to the present paper, are contributions to this study.

I

NOTE ON THE CHARACTERISTIC FREQUENCIES OF SPEECH

Speech is, in itself, a sound wave—a succession of condensations and rarefactions in the air. For the purposes of this study we are not primarily concerned with the mechanism of production, nor with the processes of perception of speech, though it may be necessary to digress to inquiries of this kind, in their bearing on certain characteristics of speech. We are interested primarily in what can be learned from the records of the speech vibrations themselves.

Speech sounds are complex, that is, they are composites of simple sounds, each component having a particular frequency, amplitude, phase and duration. Considering speech in the mass, we find its energy distributed among frequencies from 75 to above 5,000 cycles with the larger part of this energy contained in the region below 1,000 cycles. This distribution is shown approximately in Fig. 1 taken from reference (8g); the limitation on these data being that the measuring apparatus was not sufficiently sensitive to measure the speech energy associated with frequencies higher than 5,000 cycles. Inasmuch as the energy of speech resides largely in the vowel sounds, the curve in Fig. 1 can also be taken as applying to the average distribution in the vowel sounds. The energy distribution diagram is of fundamental importance in the physical study of speech sounds; it reveals at once the frequencies of large energy content which are characteristic. For each vowel sound, there is a distinctive energy frequency diagram.

The consonant sounds present a difficult problem because of the small amount of energy associated with them. Most of our knowledge of the consonant sounds is qualitative: for example Fletcher (reference 8h) who studied the nature of speech by the method of testing articulation when different frequency ranges are eliminated shows that for two fricative or sibilant consonants *s* and *z*, there are essential frequency components which lie above 5,000 cycles. The characteristic frequencies of the consonant sounds are usually only part of the whole story; these sounds are richer in transients, and clearly less periodic in their nature than the vowel sounds. And in between the two broad classes of consonant and vowel sounds there is a group of semi-vowel sounds (*r*, *l*, *m*, *n*, *ng*) closely related to the vowel group, and yielding readily a determination of their "characteristic frequencies."

There are two physical theories of vowel production; and these two theories suggest different methods of analyzing the vowel sounds into components of simpler nature. These two points of view we shall briefly consider along historical lines. We are indebted to Helmholtz for the greatest single contribution to the study of the vowels, in that he gave a complete diagram of the characteristic frequencies of the vowels (ref. 2, pp. 103–109), which was based on his celebrated experiments in analysis and synthesis by means of the Helmholtz resonators. But in connection with his scheme of characteristic frequencies he took up the theory of Wheatstone (1837) that these frequencies are true harmonic components of the cord tones, which were reenforced by resonance in the oral cavities. Some later physicists have followed this so-called *harmonic or steady state theory* of the vowel sounds, notably Miller (reference 3, pp. 239–243) who

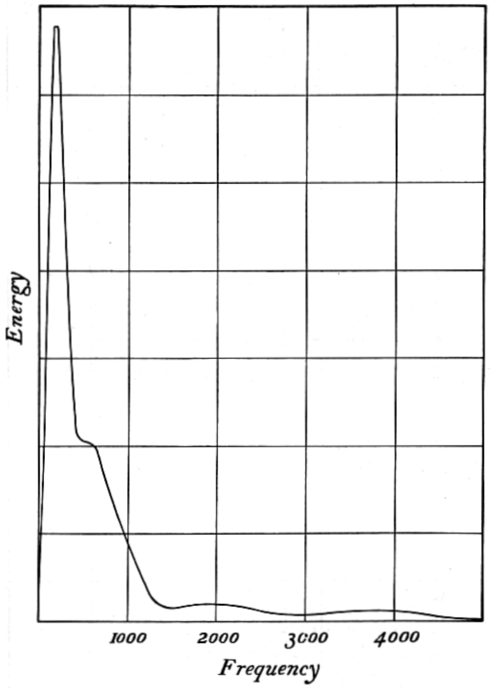


Fig. 1—Energy distribution; composite curve for male and female voices

made a very careful study of the whole matter. According to this theory the obvious procedure is to apply the classical Fourier analysis to determine the characteristic components of the vowel sounds.

Turning now to the other (and earlier) view, the so-called "Inharmonic Theory" of Willis (1829) later developed by Hermann and rather recently by Scripture (ref. (6)) we are invited to believe that the "characteristic frequencies" of the vowel sounds are the natural vibrations or *transients* in the oral cavities, when excited impulsively by the (more or less) periodic puffs of air from the glottis. According to this theory no harmonic relations need obtain between the characteristic frequencies of the vowels and the fundamental or cord tone accompanying them; and the classical Fourier analysis is not considered applicable in resolving the vowel sound into simpler components. According to this "inharmonic" or "transient" theory we must treat the natural vibrations of the oral cavities as damped vibrations and find the frequencies and damping constants of their components, as best we can from the record of the complete sound vibration.

In favor of the Helmholtz or "Harmonic" theory we have the careful studies by Helmholtz and his successors of the relations between the cord or fundamental tone, its harmonics as reenforced by the oral cavities or other resonators, and the observed characteristic frequencies of the vowel sounds. The oral cavities constitute a vibrating system of two or three degrees of freedom, the theory of which has been fully developed by Rayleigh and others, and it is to be expected that, with the speaking mechanism in normal adjustment the vowel qualities can be well accounted for by postulating harmonic forced vibrations in these cavities. This expectation has been realized in the numerous successful attempts which have been made to produce vowels artificially by using a harmonic series of tones, and reenforcing certain harmonics by suitable resonators. Miller's experiments with organ pipes (ref. 3, pp. 246–250), in which he successfully reproduced certain vowel sounds, are well known.

The Willis-Hermann theory has also suggested much notable experimental work. Scripture (ref. 6, p. 114) constructed a "vowel-organ" in which a reed pipe was used to excite the natural vibrations in resonators designed to imitate the conditions in the oral cavities, and attained some success in reproducing vowels. More recently J. Q. Stewart (ref. 8i) has produced an "Electrical Analogue" of the vocal organs with which remarkable results in reproducing vowel sounds and even some of the consonant sounds have been obtained. In this electrical arrangement transients excited by an interrupter in oscillatory circuits take the place of the transient vibrations of the oral cavities. Finally Paget (reference (9a) below) has constructed a whole series of double resonators which may be excited by blowing air into them through an "artificial larynx," and from which he has obtained all of the vowel sounds. As the result of this work he has given a very complete chart of the characteristic frequencies of the vowels and he has been led to the conclusion that there are *two* characteristic frequencies or regions of resonance for each vowel sound.

From the standpoint of practical acoustics both theories have contributed to progress, and it seems that the experimental physicist would not be justified in partiality to either view. Speech is a variable phenomenon; the cord tones are not always stable; in speaking and in singing there are allowable variations in duration, intensity and frequency of the component tones without essential change in the characteristics of the vowel sounds. Given accurate records of the speech sounds as normally pronounced by a number of speakers, we should expect to arrive at nearly the same characteristic frequencies whichever mode of analysis we adopt. As pointed out by J. Q. Stewart (Ref. 8i) Rayleigh has stated (Sound, Vol. II, p. 473) that the disagreement between the Helmholtz-Miller, or steady state theory of vowels, and the Willis-Hermann-Scripture, or transient theory is only apparent; to quote Stewart, "The disagreement concerns methods rather than facts. Which viewpoint should be adopted is thus a matter of convenience in a given case. When the transmission of speech over telephone circuits is in question, for example, the steady state theory often possesses obvious mathematical advantages. On the other hand, the quantitative data relating to the physical nature of vowels which are given in D. C. Miller's well-known book "The Science of Musical Sounds" expressed as they are in terms of the steady state theory are less compact and definite than the data of Table I (Stewart's paper) which are expressed in terms of the transient theory. The general agreement between the two sets of data is, of course, obvious."

In studying the behavior of vibrating systems from the theoretical standpoint, there is a tendency to emphasize the intimate relations that exist between transient and steady state phenomena. Both depend only on the driving forces and the constants of the system,

---

[9](a) Sir R. A. S. Paget: "The Production of Artificial Vowel Sounds." Proc. Roy. Soc. A102, Mar. 1, 1923, p. 752.

[9](b) A second memoir: "The Nature and Artificial Production of Consonant Sounds." Proc. Roy. Soc. A 106, Aug. 1, 1924, p. 150, to which reference will be made in more detail later.

Other papers by Paget include: Nature, Jan. 6, 1923, "Nature and Reproduction of Speech Sounds." Electrician, Apr. 11, 1924. The Same Title. Proc. Land. Phys. Soc. 36 pt. 3, Apr. 15, 1924, p. 213: Discussion on Loud Speakers.

hence "the solution for transient oscillations of the system is reduced to formulae which are functionally the same as those for steady state oscillations" (reference 10; see also reference 11). But before leaving this discussion of speech characteristics it should be noted that the essence of the matter lies not so much in reconciling the two theories of the vowel sounds as in ascertaining what motions really take place in the oral cavities, and in the air near the vocal cords. Though the process of harmonic analysis is to be applied to the records of the vowel sounds, we must recognize its limitations, and not necessarily infer steady state conditions. Indeed the most casual inspection of the records shows a certain *lack* of periodicity in the phenomena recorded; and it is hardly to be expected that all the phenomena can be satisfactorily summed up on the basis of the harmonic theory.

II

THE RECORDING APPARATUS[12]

In providing means for accurately recording sound waves, use has been made of three devices recently developed in this Laboratory and we believe that by properly connecting these together we have obtained a recording instrument which is superior in accuracy and power to any heretofore used. These three devices were each nearly free from distortion, and such residual distortions as could not be eliminated were so controlled that they practically offset one another over a wide range of frequencies.

The first element in the recording set is the condenser transmitter, which has been thoroughly investigated by Wente (refs. 8b, 8c, 8f); its frequency characteristics, in both amplitude and phase are shown in Fig. 2. The particular transmitter used was of recent design and had been carefully standardized and calibrated especially for this work.

The condenser transmitter was connected to the input terminals of a seven-stage amplifier as shown in the large diagram of Fig. 5 which gives the details of the electrical circuit, including the third

---

[10] J. R. Carson: Phys. Rev. X, 1917, p. 217, "On a General Expansion Theorem for the Transient Oscillations of a Connected System."

[11] T. C. Fry, Phys. Rev. XIV, 1919, p. 117. "The Solution of Circuit Problems."

[12] Thanks are due to Messrs. C. F. Sacia and C. J. Beck for the skill and care with which they assembled and calibrated the recording apparatus, and made the complete set of records. The writer is also under obligation to Mr. Sacia for aid in choosing the sounds to be recorded, and systematizing the collection; Mr. Sacia also developed and applied the photomechanical method of analyzing records, the results of which are given in Figs. 13 and 14 of this paper.

element, a special oscillograph, which was connected to the output terminals of the amplifier. The first six tubes, in cascade, provided a voltage amplification of about 40,000; the last eight tubes, in parallel, constituted a "current transformer" working into the low impedance of the oscillograph vibrator, with a small resistance in series. The coupling between the stages, and between amplifier and terminal apparatus, was entirely of resistance and capacity, with the capacity reactance minimized. In all tests of the circuit the condenser transmitter and the oscillograph vibrator remained in their fixed positions, as shown in the diagram, so as not to disturb the electrical characteristics of the circuit. The frequency characteristics of the amplifier in amplitude and phase are shown in Fig. 3. In measuring the amplitude characteristic a small electromotive force was introduced in series with the transmitter, in the input mesh; and in measuring the phase lead of the output as a function of frequency use was made of the Alternating Current Potentiometer of Wente (Jour. A. I. E. E. Dec. 1921) the other details of procedure being as usual.

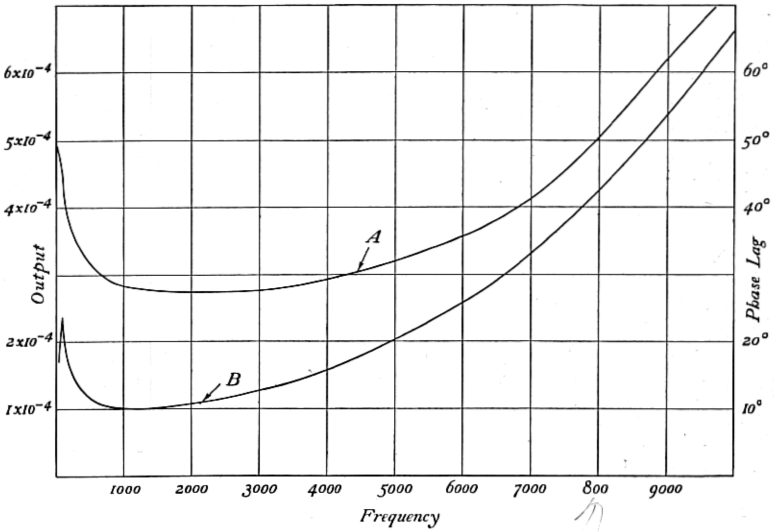


Fig. 2—Curve *A*: Output of transmitter in volts per dyne per sq. cm. Curve *B*: Phase lag of voltage behind pressure in condenser transmitter

The characteristics of the oscillograph vibrator are shown in Fig. 4. This vibrator was specially constructed, with small mass, high tension and damping; when the requisite dynamical characteristics were once obtained, its calibration presented no great difficulty.

In combining the transmitter, the amplifier and the oscillograph to form the complete recording apparatus there were two primary requirements; first, the set as a whole should be free from frequency distortion in both amplitude and phase, and second, the output of the set as a whole should be a linear function of the input within the working energy range at each frequency. The first of these conditions is in

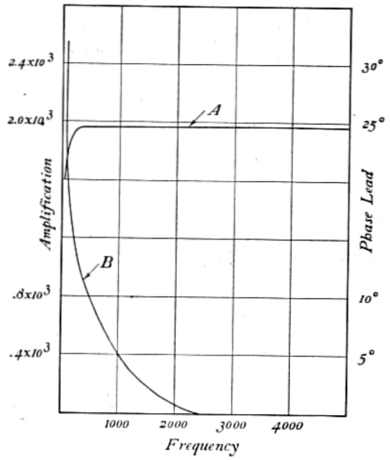


Fig. 3—Curve *A:* Amplitude frequency characteristic of amplifier. Curve *B:* Phase lead of output, vs. frequency of voltage input to amplifier

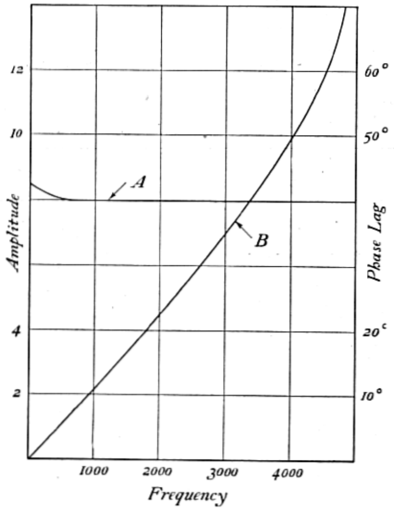


Fig. 4—Curve *A:* Amplitude frequency characteristic of oscillograph. Curve *B:* Phase lag of amplitude behind current in oscillograph

general the harder to fulfil. Frequency-amplitude distortion has been practically eliminated as we have seen from each of the three essential parts of this apparatus; and although it was found impracticable to make each part of the apparatus free from frequency distortion in phase, it was possible to give the complete set good frequency characteristics in both amplitude and phase as will be explained.

In a vibrating system of one degree of freedom when we wish to avoid frequency distortion in amplitude, we usually adjust the resonant frequency so that it is above the range of frequencies within which we desire to work; in addition, it is desirable in most cases to make the damping of the system large. With these adjustments made it is found that there is a phase lag between amplitude and driving force which rises with frequency and reaches a maximum above the resonant frequency, and it is possible to make this phase lag nearly proportional to the frequency over the range of frequencies within which it is desired to work.

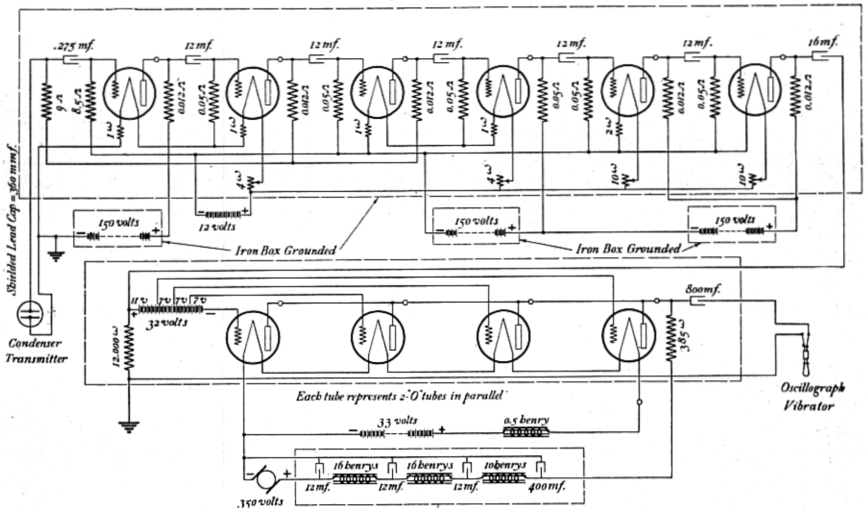


Fig. 5—General diagram of recording apparatus showing circuit details

It is well known that if equal driving forces produce equal amplitudes at all frequencies, and if the phase lag of the amplitude with respect to the driving force is proportional to frequency, then a driving force of complex wave form is reproduced without distortion of wave form in the vibrating system. These conditions held very well over the desired range of frequencies in the oscillograph vibrator, as shown in Fig. 4. In the case of the condenser transmitter, however, there were departures from these conditions in the frequency interval from zero to 500 cycles for which allowance had to be made.

In the amplifier the effect of capacity reactance was nearly eliminated. Owing to the small remaining capacity reactance there was a phase lead of amplifier current with respect to driving force which was applied to offset the excessive phase lag in the condenser transmitter at the low frequencies. The particular adjustment of amplifier finally arrived at represented the best compromise, considering the difficulty encountered with the transmitter characteristics. With this compromise made there was an unavoidable phase lead in the whole apparatus for frequencies below 125 cycles, but this was not serious as most of the speech energy is in higher frequencies. After all final adjustments were made the overall frequency characteristics of amplitude and phase were as shown in Fig. 6. Thus ultimately there was obtained a system with practically uniform amplitude characteristic from 500 to 5,000 cycles, without serious departure from this level for frequencies from 50 to 500 cycles; and with phase lag nearly a linear function of frequency from 125 to 5,000 cycles, after passing through a period of lead in the narrow interval from 50 to 215 cycles.

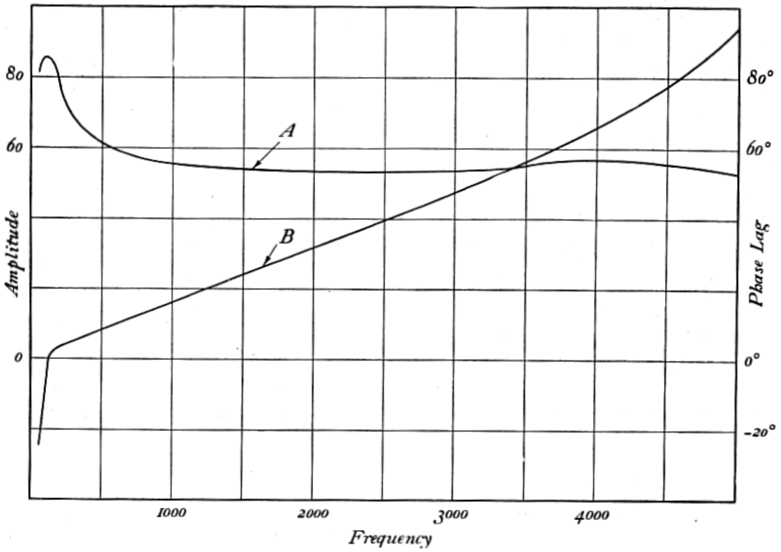


Fig. 6—Overall frequency characteristics of amplitude and phase of the recording system. Curve *A:* Oscillographic amplitude per unit of pressure on transmitter diaphragm. Curve *B:* Phase lag of oscillographic amplitude behind pressure on diaphragm

Consider now the second requirement which the recording system had to meet: namely, that the output of the system should be a linear function of the input within the working energy range at each frequency. Thorough investigation of the condenser transmitter had shown that this instrument met this second requirement very well; it was only necessary to test the remainder of the system. Fig. 7 gives the results of these tests, the voltages introduced in series with the transmitter at the input being maintained at different constant levels, while the frequency was varied. An inspection of the data shows that this requirement was very accurately fulfilled, by the whole electrical system.

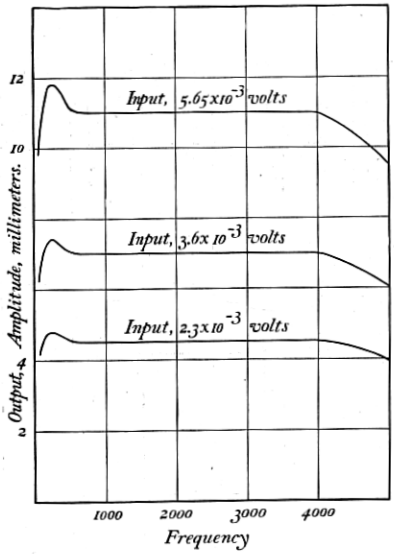


Fig. 7—Amplitude frequency characteristics of circuit-oscillograph at different energy levels

Returning now to the overall characteristics of the apparatus, it was thought advisable to test the calibrations in amplitude and phase lag by comparing the computed and the observed distortion when a square-topped acoustic wave was impressed on the apparatus. The steep sides and the flat tops of these waves can be reproduced without distortion only if the apparatus possesses first class characteristics, both in amplitude and phase lag, and the test was a severe one. As would be expected from the calibration curves of Fig. 6 there was a certain amount of distortion in recording this wave, and the square-topped wave, with its very large fundamental component, made this distortion appear much worse than would an ordinary speech wave.

Fig. 8 illustrates the apparatus used to produce the acoustic square-topped wave. An electrode resembling the back plate of the condenser transmitter was mounted in front of the transmitter diaphragm. Between this electrode and the diaphragm was applied a high potential which was made alternately positive and negative by a commutator.

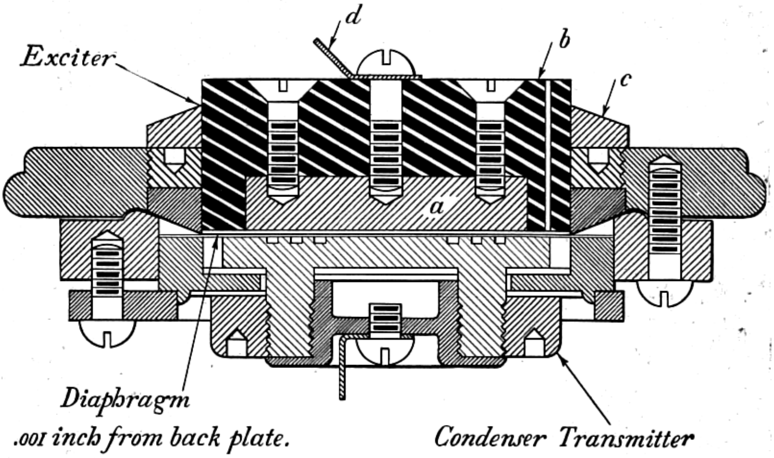


Fig. 8—Condenser transmitter coupled with square-topped-wave exciter

Exciter Parts

a. Steel Electrode 0.006 inch from Diaphragm. *b*. Micarta Insulation.
c. Supporting Ring. *d*. Electrode Terminal.

By this arrangement the desired positive and negative pressures were produced on the diaphragm. The distance between the auxiliary electrode and the transmitter diaphragm was about .006 inch. This electrostatic coupling was found to be sufficiently close to give a suitable deflection of the transmitter diaphragm, while the stiffness and damping of the air film did not alter the dynamical characteristics of the transmitter.

Fig. 9 is an oscillogram showing the wave form recorded by the apparatus when acoustic square-topped waves of frequencies 84, 153 and 306 cycles per second are impressed on the transmitter. Timing waves of frequencies 75, 150 and 300 are also shown. Analysing the original wave by the Fourier method, and allowing for the distortion in amplitude and phase of each component frequency, a computation has been made of the wave form in the output in the case of the square-topped waves of 84 and 153 frequency. The results are shown in Fig. 10, The Fourier series representing the 84-cycle wave contained 30 terms, the component frequencies being odd multiples of 84 up to a limit of 4,956 cycles; for the series representing the 153-cycle wave 17 terms were used covering the range from 153 to 5,049 cycles. The agreement between calculated and observed output waves would have been more exact, particularly at the corners of the wave shapes, if calibrations and calculations had been carried to frequencies considerably above 5,000. As it was, the performance was considered good; it indicated that the uncorrected records of speech waves as taken were sufficiently accurate for most purposes, while if harmonic analysis of the records was planned accurate results could be obtained over the range from 80 to 5,000 cycles, if the correction factors determined by the calibration were applied.

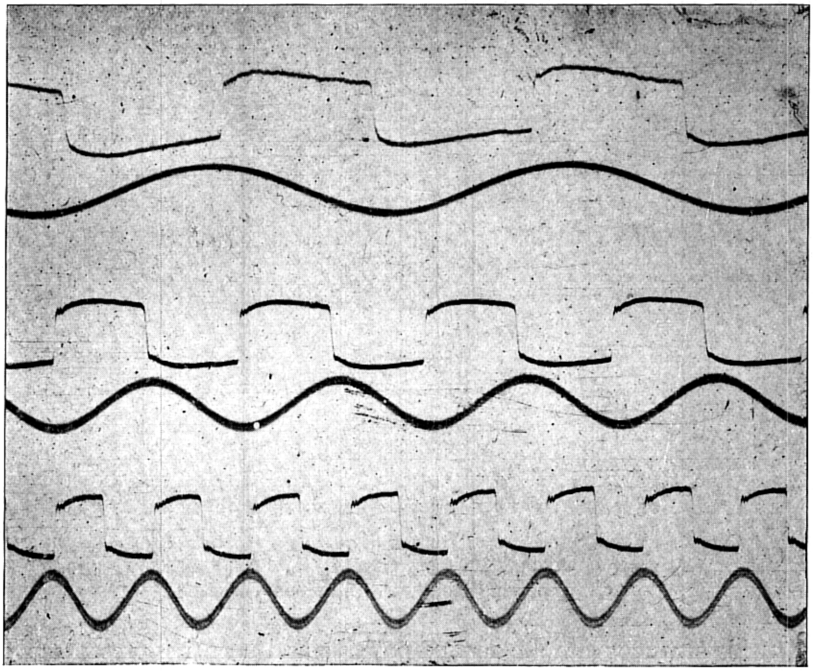


Fig. 9—Oscillogram of square-topped acoustic waves as recorded by the apparatus

In this description of the recording apparatus the emphasis has been placed on the dynamical characteristics of the apparatus and its calibration, but some of its other working features may briefly be mentioned. The apparatus was sufficiently powerful to record sounds spoken in an ordinary tone of voice, with the speaker's mouth about three inches from the transmitter. A key was pressed by the speaker just before the sound was spoken, this releasing a shutter placed before a rotating film drum on which the record from the oscillograph vibrator was traced. The film drum was some 52 inches in circumference, and there was mounted on it a length of Eastman super-speed film with which records could be made at a peripheral speed of about 20 feet per second. Thus each hundredth of a second corresponds to two inches or more in the time scale on the film. Besides opening the shutter, the key released a mechanism which swung the oscillograph vibrator through an arc during the progress of the record, thus tracing a helical record on the film. By this means records up to 200 inches in length, or for nearly one second of duration were taken. The average length of the wave trains recorded was less than 0.5 second; thus it was possible to graph the pressure wave of the whole speech sound from beginning to end. Immediately following the recording of the speech sound a timing wave of 1,000 cycle alternating current, taken from a standard oscillator, was recorded on the film at one side of the speech record, without disturbing the speed adjustment of

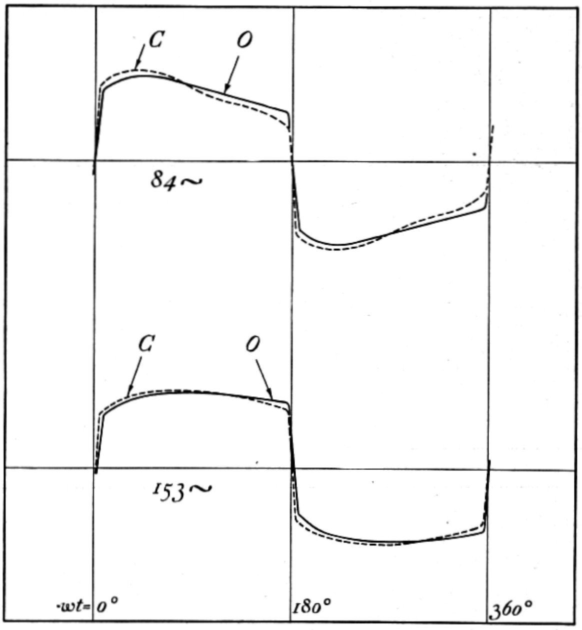


Fig. 10—Calculated and observed wave forms, as recorded by the apparatus

the rotating drum. Thus the time scale was accurately determined for each record.

Especial care was taken with the optical system to insure fine definition and strong illumination of the spot on the film and the films were developed for maximum contrast. As a result, the records were sufficiently clear to permit their reproduction by the line-engraving process. Each of the plates shown in this paper is made up of overlapping sections from the original record, each faithfully reproduced, and the whole arranged to give the complete record within the limits of one page. A section of one of the original records as taken is shown in the figure above.

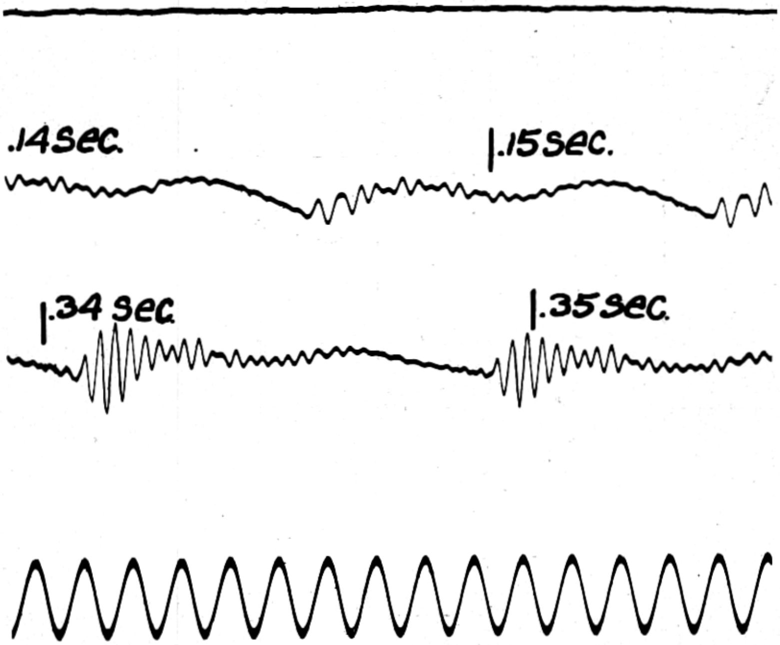


Fig. 11—Section of original record showing timing wave

III

Classification of the Records

In selecting and classifying the vowel sounds for record, use has been made, with slight alteration, of the phonetic arrangement adopted by Fletcher (ref. 8 h). This arrangement of the vowel sounds is illustrated in the diagram of Fig. 12. In this diagram eleven standard "pure-vowel" sounds from *oo* to long *e* are arranged according to the conventional "triangle" and two related vowel sounds *ar* and *er* are interpolated in their proper places. A group of eight records was made of each of these thirteen vowel sounds, four in each group by male voices, and the other four by female voices. Each of these records, Plates 1 to 104 (Groups I to XIII), represents the vowel sound as spoken naturally, and continuously recorded from beginning to end.

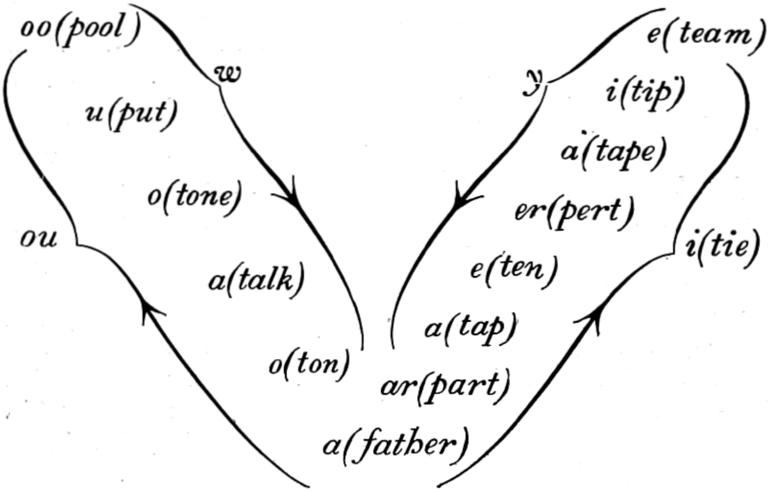


Fig. 12—Classification of vowel sounds

No attempt was made to record the vowels *w*, *y*, *ou* and long *i*. These usually have transitional characteristics which are sufficiently indicated by the arrows in the diagram. The first two of these, when followed by vowels, and the last two, in nearly all cases, fall into the class of diphthongs.

Following the groups of records of the "pure-vowel" sounds of the diagram it was originally planned to make a group of records of the semi-vowels *l*, *m*, *n*, *ng*, and *r*, recorded in connection with certain vowels. It seemed best however to present records for the sounds *ar* and *er* in connection with the standard vowel sounds as noted above (*ar*, *er*, Groups VII, X) and only these records of the sound *r* were taken. The four remaining sounds were arbitrarily divided into two groups because of the number of records made, and the first of these (Group XIV) contains records of *l* and *ng*. These were made by two male speakers, using the syllables *loo*, *lee*, *la* and *ngoo*, *ngee*, *nga*.

Group XV is devoted to the semivowels *n* and *m*, each recorded with the three vowel sounds *oo*, long *e* and *a*, by the two male speakers, as in the preceding group. Groups XIV and XV are intimately related, and as will appear the four semi-vowel sounds are closely related to the vowel diagram.

When this study was planned, it was thought that the apparatus would be particularly adapted to recording vowel sounds and no great hopes were entertained of applying it to definitive investigation of the consonant sounds. As the work progressed however, it was found that some of the characteristics of the consonant sounds could be recorded and the program was enlarged to include the records of Groups XIV to XVII inclusive. Each of the records of a consonant and vowel combination can be compared with the corresponding record, by the same speaker, of the pure vowel alone in one of the earlier groups, and certain conclusions as to the nature of the consonant sound can be formed.

Group XVI includes records of the six stop (or "hard") consonants *b*, *p*; *d*, *t*; *g*, *k*; followed by two transitional consonants *dth* (as in *then*) *th* (as in *thin*); each associated with the vowel *a*, and recorded by the two male speakers. The natural arrangement is in pairs, the related voiced and unvoiced variations being grouped together.

The last Group (XVII) includes records of eight fricative ("soft" or "sibilant") consonants paired in the same way. These are *v*, *f*; *j*, *ch*; *z*, *s*; *zh* (azure), *sh*; each associated with *a* and recorded by the two male speakers.

The following table lists in groups all the records made. As it is not practicable to engrave and print with this article the whole set of

TABLE I

Complete List of Speech Records

Group			Plates
I	*oo* as in *pool*,	by Eight Speakers	1– 8
II	*u* as in *put*,	by Eight Speakers	9– 16
III	*o* as in *tone*,	by Eight Speakers	17– 24
IV	*a* as in *talk*,	by Eight Speakers	25– 32
V	*o* as in *ton*,	by Eight Speakers	33– 40
VI	*a* as in *father*,	by Eight Speakers	41– 48
VII	*ar* as in *part*,	by Eight Speakers	49– 56
VIII	*a* as in *tap*,	by Eight Speakers	57– 64
IX	*e* as in *ten*,	by Eight Speakers	65– 72
X	*er* as in *pert*,	by Eight Speakers	73– 80
XI	*a* as in *tape*,	by Eight Speakers	81– 88
XII	*i* as in *tip*,	by Eight Speakers	89– 96
XIII	*e* as in *team*,	by Eight Speakers	97–104
XIV	Semi-Vowels *l*, *ng*	by two male speakers	105–116
XV	Semi-Vowels *n*, *m*	by two male speakers	117–128
XVI	Six Stop Consonants; transitional *dth*, *th*; by two male speakers		129–140
XVII	Eight Fricative Consonants, by two male speakers		145–164

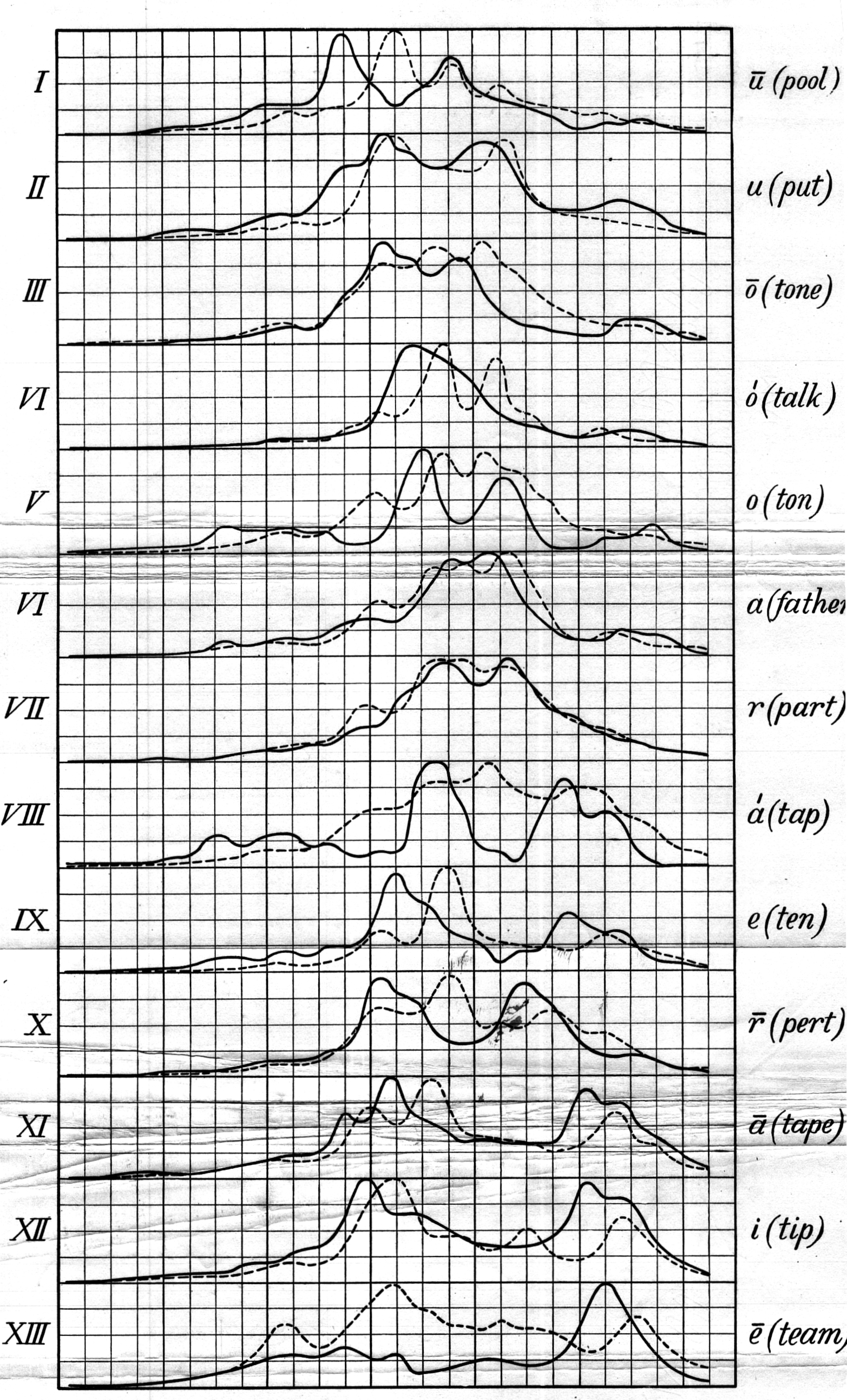


Fig. 13—Analyses of vowel sounds. Relative importance of the amplitudes at different frequencies taking into account the sensitiveness of the ear

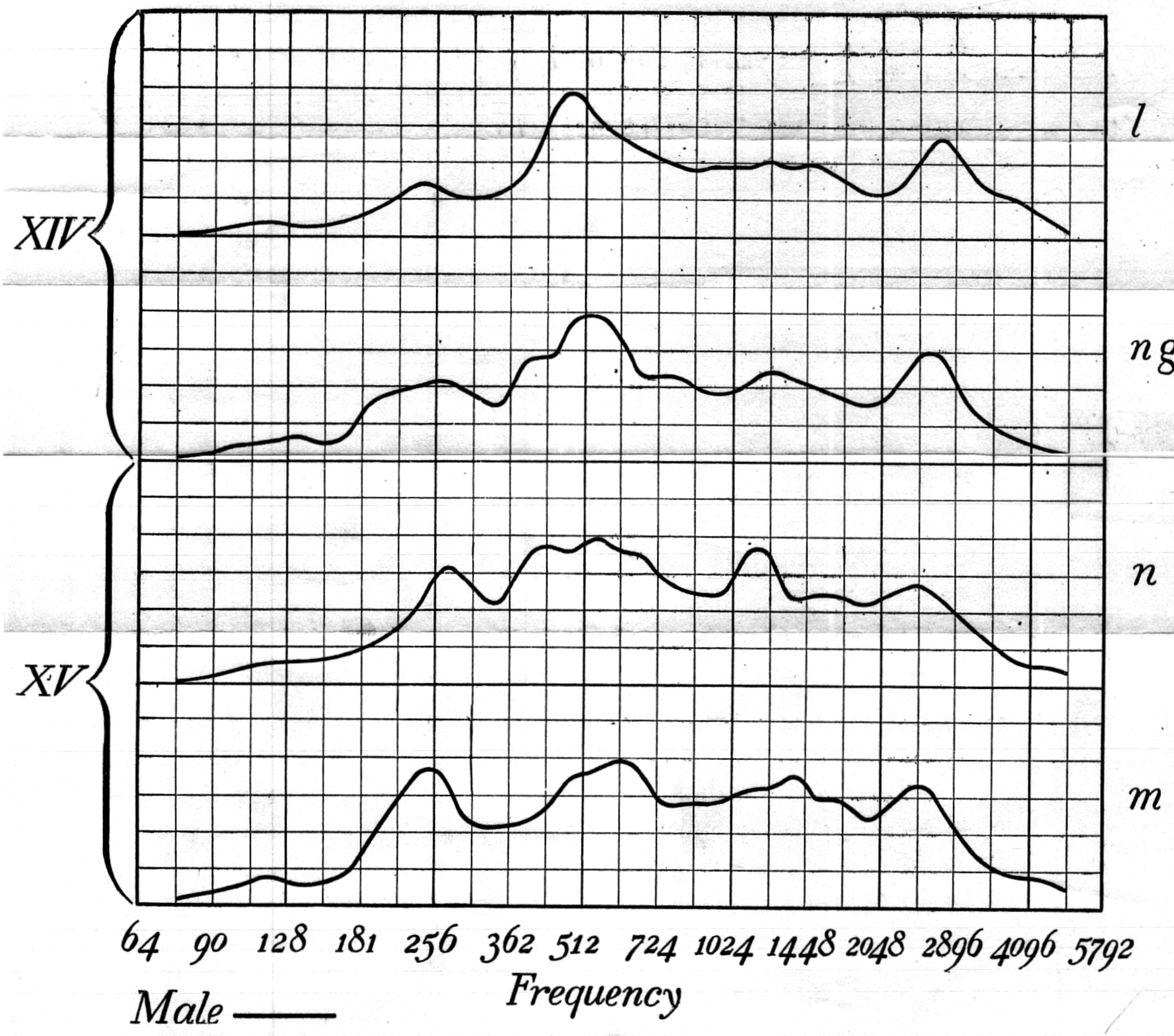


Fig. 14—Analysis of four semi-vowel sounds. Relative importance of the amplitudes at different frequencies taking into account the sensitiveness of the ear

160 records, a selection has been made of some 13 typical examples which illustrate characteristic consonant and vowel wave forms. These are listed in table II and their properties are described in detail in the following sections. It may not be amiss to summarize here the basis on which these particular records were chosen for publication.

TABLE II

List of Records Shown in This Paper

Record No.	Plate No.	Title	Speaker
143	9	*u* as in *put*	MA
192	40	*o* as in *ton*	FD
139	41	*a* as in *father*	MA
151	49	*ar* as in *part*	MA
148	89	*i* as in *tip*	MA
234	108	*lee*	MB
238	110	*la*	MB
229	124	*moo*	MB
286	136	*ta*	MB
289	138	*ga*	MB
272	151	*cha*	MA
293	158	*za*	MB
294	160	*sa*	MB

The most important sound (*a*, as in father) is represented in 7 of these records, which include six instances of its combination with other sounds. The record of *ar* (Plate 49) which was chosen is the most characteristic and interesting one of its group. The other vowel records (Plates 9, 40, 89) are sufficiently scattered about the vowel triangle to give an idea of the variation in the high frequency characteristics which is to be an important subject of discussion later. One record of a female voice (Plate 40) is probably sufficient to show the distinctive fundamental, about an octave higher, characteristic of such records. Plate 108 was chosen to show the resemblance between l and e, which establishes a natural transition between the vowel and semi-vowel sounds. From plates 108, 110 and 124 a good idea of the relative amplitudes of vowel and semi-vowel sounds can be obtained; a similar observation holds in the comparison of the vowel and consonant sounds of Plates 136, 138, 151, 158 and 160. Plates 136 and 138 show two extended transients of moderate frequency, the latter in connection with a voiced consonant (*hard g*); Plate 151 is similar to 136—but the vowel following the consonant is less suddenly produced. The pair, Plates 158 and 160, show the voiced and unvoiced hiss (z and s respectively) a sound of very high frequency, which is the limiting case of this type of consonant.

The plates reproduced with this paper are reduced slightly (15 or 20 per cent) in scale, as compared with the original records, to bring them within the page height of the Journal.

In producing this system of records we believe that we have covered the speech sounds as fully as we are justified in doing with the present recording apparatus. In the case of each vowel the combined data from the eight records constitute a sufficient basis for the most thorough harmonic analyses that can be made and they should yield accurate results for the characteristic vowel frequencies. In analysing these records small corrections are of course necessary on account of the slightly imperfect frequency characteristics of the apparatus, but these corrections can be taken without difficulty from the calibration curves.

The amplitude scale in these records is arbitrary in each case. This is for the reason that, owing to the widely different conditions of voice control among the different speakers, the recording apparatus had to be adjusted to different levels of sensitiveness for each record in order to obtain the requisite maximum oscillation of from 1 to 2 centimeters. No attempt has been made to compare the absolute amplitudes from one record to another on account of these intensity variations. The emphasis has been placed rather on obtaining in each record a good well-defined wave which could be enlarged if necessary.

Notwithstanding the fact that for frequencies above 5,000 cycles the apparatus was not nearly as good as for frequencies within the calibration range from 75 to 5,000 cycles, the records obtained of some of the consonant sounds are of considerable practical value. It is felt however, that the present apparatus has been used nearly to the limit of its possibilities and that devices other than the usual oscillograph vibrator offer more promise in any further investigation of the consonant sounds. It is planned later to issue a more complete set of these records as a supplement to the present paper in order to make the collection available to those especially interested.

IV

STATISTICAL STUDY AND HARMONIC ANALYSIS OF THE VOWEL SOUNDS

A detailed inspection of the records taken, and particularly of the records of the vowel groups shows that much labor would be required to analyze these records throughout their length, according to the usual methods of harmonic analysis. In nearly every case it would be impossible to obtain the mean energy distribution in a given record, allowing for variations from cycle to cycle of the fundamental, by choosing from each record only a few such cycles as representative and analyzing these.[1] If, for example, only 10 cycles were taken at selected intervals from each of the 104 vowel records shown there would be required over one thousand such analyses, and to be of value these analyses should include components of frequency from 100 to 5,000 cycles. For this reason a mechanical method of analysis has been applied to determine from the records the average frequency spectra of each of the vowel and semi-vowel sounds.

First let us consider the vowel records in a simpler and more general way. Considerable information has been obtained by inspection, using such simple apparatus as a pair of compasses and a rule in connection with the time scale on the records. The time scale greatly facilitates the process; it is in most cases possible to count the number of cycles of any one prominent component occurring in an interval of .01 second, and by doing this in various parts of the record, to arrive at a rough average frequency for the component in question.

In the case of the low frequency components (the fundamental and the lower characteristic frequency) the procedure was to make this examination at 3 points; one near the start, one near the middle, and one near the end of each record. In this way the most significant changes in pitch and wave form during the course of the record can be brought to light, and some of the individual characteristics of the speaker revealed. A statistical compilation of these results serves to show certain "normal" characteristics of pitch variation, and permit the detection of a certain amount of "personal bias" of the individual speaker in his departure therefrom. In the examination of the low frequency characteristics a note was made as to the harmonic relation between the fundamental and the lower characteristic frequency; of the amplitude of the lower characteristic frequency as being greater or less than the amplitude of the fundamental; and of the behavior of the amplitude of the lower characteristic, during the cycle of the fundamental. The amplitude of the low frequency characteristic is either substantially constant during the cycle or falls away as a transient vibration.

The high frequency components are clearly shown in the records, but it is more difficult to determine their exact frequencies, and practically impossible to relate them harmonically to the fundamental. These oscillations were counted in from four to eight locations in each

---

[1] It is practicable, however, to obtain valuable data as to the formation of the vowel sounds by analyzing separately the successive cycles at the beginning of a typical vowel record. A study of this kind, based on these records, is being carried out by Messrs. N. R. French and W. Koenig of the American Telephone and Telegraph Company.

record, and a maximum and minimum figure determined for the frequency wherever possible. The behavior of the amplitude of the high frequency component during the cycle was noted, and a rough estimate made of its magnitude. Practically all the vowel records show frequencies above 2500 cycles and the amplitudes in some cases are large. In only two records out of 104 was the high frequency component too small in amplitude to give a frequency determination. These high frequency components may or may not be characteristic of the given sound; this question is more fully dealt with later.

To complete the examination of each record its duration was noted, and this time was divided into three intervals: (1) a building up period in which the oscillations rise from zero to an amplitude which shows all the components clearly; (2) a middle period in which the general amplitude remains nearly constant, but in which some variations in the amplitudes and phases of the component frequencies usually take place; and (3) a period of decay in which the components disappear and the oscillation gradually loses its characteristic wave form.

The procedure may be illustrated by its application to the first record for which the following data were recorded:

Plate No. 1, *oo* as in *pool*. Speaker MA. (Male).

Time to build up, .05 sec.; Middle period, .20 sec.; Period of decay, .06 sec.; Total Duration .31 sec.

Fundamental: 102 at start, rises to 108 in middle, rises to 120 at end. Pitch Variation normal. (See explanation below).

Low Frequency Characteristic: 400 at start, 430 at middle, 440 at end. Amplitude greater than that of fundamental. Approximately, a fourth harmonic of fundamental, but amplitude variation during the cycle suggests a transient.

High Frequency Component: Minimum, 3300 cycles. Maximum, 3600 cycles. Noticeable throughout; amplitude variation suggests a transient.

No other frequencies.

This routine was applied to each of the 104 vowel records and a general summary made of the results, giving approximate values of the vowel characteristics which forecasted the more accurate results obtained later from the mechanical harmonic analysis.

The simplest phenomena to summarize are the general characteristics of the individual speakers. These are based on the mean performance of each in speaking the thirteen vowel sounds, and will be useful in the discussion to follow; they are shown in Table III, below:

TABLE III

Speakers' Characteristics

Male Speakers	Mean Fundamental Pitch at Start, Middle and End	Mean Pitch	Mean Duration of Records
MA—low pitched	97–105–111 (normal)	104	.275 sec.
MB—low pitched	112–115–112 (biased)	113	.222 biased toward short records
MC—high pitched	124–131–134 (normal)	130	.235 (biased toward short records)
MD—high pitched	134–148–175 (normal)	152	.305
	Mean for male Speakers	125	.259 sec.
Female Speakers			
FA—low pitched	224–241–209 (normal)	224	.290 sec.
FB—low pitched	256–251–194 (biased)	234	.373 biased toward long records
FC—medium	233–255–244 (normal)	244	.320
FD—high pitched	271–274–279 (biased)	275	.348 (biased toward long records)
	Mean for female speakers	244	.333 sec.
	Mean duration		.296 sec.

These records were made without constraint imposed on the speaker, except that he had to start and stop within an interval of about one second, and was requested to repeat the sound several times at what he judged to be constant loudness. The resulting variation in performance may therefore be of some interest.

Of 52 men's records the vowel sounds 35 records showed a "normal" effect of progressive *rise in pitch* during the course of the record. (The mode is taken as the normal effect, and follows the mean very closely.) In 6 records out of 13, speaker MB showed an individual or biased effect of slight fall in pitch toward the end. The women's records show greater variation, 24 records out of 52 showing a "normal" effect of a *rise in pitch, followed by falling pitch*, during the course of the record. The individual bias of speaker FB toward progressive fall in pitch was shown in 7 records; that of FD toward progressive rise in 4 records.

The relative constancy in fundamental pitch shown by speaker MB is best exemplified in Plate No. 58. Speaker FD made 3 records of constant pitch: Nos. 24, 40 and 48. Other records of constant pitch are Nos. 19 and 99, both by MC.

In duration, the bias of speaker MB towards short records was shown in 6 records which fell short by .08 sec. or more of the mean for the particular sound considered; that of MC also in 6 records according to the same test. Speaker FB produced 5 records, and speaker FD, 2 records too long by the same amount.

Consider now the general properties of the spoken vowel sound, as deduced from these records. First there is a period of rapid growth in amplitude, lasting about 0.04 second, during which all components are quickly produced, and rise nearly to maximum amplitude; second the middle period, the characteristics of which have been noted, lasting about 0.165 second, followed by the period of gradual decay lasting about 0.09 second, bringing the total length to approximately 0.295 second. There is a tendency to short duration among the "short" vowels (eg. short *o*, *e*, *i*) and a tendency to longer records among the broader sounds, as might be expected.

The behavior of the fundamental frequency (or "cord tone") during the course of the record will follow normal or individual characteristics as has been described.

The low frequency characteristic appears early, usually before the fourth cycle (for men) or before the seventh (for women) and normally is in harmonic relation with the fundamental. In the eleven pure vowel sounds (omitting the *ar* and *er* groups) this point was examined at 264 locations in 88 records with the result that the harmonic relation obtained in at least 214 cases. On the other hand the normal behavior of the amplitude of the low frequency characteristic suggests the decay of a transient oscillation during each fundamental cycle—this effect being noticeable in at least 64 of the 88 pure vowel records. This transient effect was also noticeable in 13 of the 16 records of *ar* and *er*, where the harmonic effect was not so noticeable. The appearance of the transient effect depends to some extent on the relative frequencies of the fundamental and the characteristic; where the fundamental period is short, (as often in the case of the women's records) there is not sufficient time for decay of the characteristic tone before it receives a new impetus in the next cycle of the fundamental.

As noted above, all the records contain high frequency vibrations which are of such amplitude that they suggest characteristic frequencies. A general mean of these frequencies would be in the neighborhood of 3200 cycles, and in the case of two records by speaker FC (Group I and Group XIII) the frequency rises to about 5000 cycles. Recalling the usual classification of the vowel sounds into two groups—(1) those of "single" resonance, placed on the left leg of the triangle, (Fig. 12) and (2) those "double" resonance placed on the right leg of the triangle—there are some differences in the behavior of the high frequency components which can be related to these broad classes.

TABLE IV

Statistical Data From 104 *Records of Vowel Sounds*

Sound	Duration				Mean Fundamental Frequency		Mean Low Characteristic Frequency		Scattered Low Freq.		Mean High Characteristic Frequency		Scattered High Freq.	
	Start	Middle	Decay	Total	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
I *oo*(pool)	.061	.164	.126	.351	140	270	411	581	750 (1)	1200 (1)			3700 (4)	4412 (4)
II *u*(p*u*t)	.057	.115	.077	.249	138	250	457	691	988 (4)	1100 (3)			3637 (4)	4250 (4)
III *o*(tone)	.053	.139	.133	.325	116	237	520	729	830 (3)	1112 (4)			3475 (4)	3700 (4)
IV *a*(talk)	.034	.191	.065	.290	112	243	722	801	950 (2)	1150 (2)			3612 (4)	4075 (4)
V *o*(ton)	.046	.179	.061	.280	118	253	654	854	1100 (4)	1188 (4)			3212 (4)	3353 (3)
VI *a*(f*a*ther)	.029	.199	.078	.306	113	234	955	1036	1150 (2)	1425 (2)			3683 (4)	4200 (3)
VII *ar*(p*ar*t)				.345	110	231	630	701	917 (3)	1012 (4)	1965	2162	3800 (2)	4150 (1)
									Note 1		Note 1			
VIII *a*(tap)	.038	.180	.076	.294	123	232	796	960			1900	2165	3150 (3)	3175 (2)
IX *e*(ten)	.034	.119	.066	.219	121	247	612	775			1800	2000	2925 (4)	2925 (4)
											Note 2		Note 2	
X *er*(pert)				.331	131	239	570	712			1688	2188	3050 (2)	3500 (1)
XI *a*(tape)	.042	.172	.091	.305	125	235	494	614			3000	2800		
XII *i*(t*i*p)	.036	.126	.049	.211	137	253	450	523			2950	2962		
XIII *e*(team)	.036	.189	.116	.341	136	252	296	332			2987	3266		4800 (1)
Means, or "Normals"	.042	.161	.085	.288(11) .296(13)	125	244								

NOTE 1—Both of these sets of frequencies must be characteristic of *ar*. (Compare Fig. 13, also the results of Paget, quoted later.)

NOTE 2—The high frequency characteristics are less definitely located, for short *e*, than for any other doubly resonant vowel sound. (Compare Fig. 13.) The two sets of frequencies given above define a band of frequencies centered about 2,400 cycles within which the characteristic high frequency must be contained.

In the sounds of the first class the high frequency component is usually small in amplitude, more subject to individual bias in its frequency, and may or may not build up in amplitude as early as the low frequency characteristic. In the sounds of the second class the high frequency characteristic is usually prominent from the start and builds up very rapidly; while there is less variation in its frequency with the individual speaker. In sounds of the first class there is no decided suggestion of a transient in the high frequency (23 out of 40 records, Groups I to V inclusive) while in sounds of the second class the transient effect is pronounced (39 out of 40 records, Groups VIII, IX, XI, XII, XIII).

With these considerations in mind there is presented in table IV a summary of the data obtained from this preliminary examination of the vowel records. The mean duration time, and its subdivisions, are shown in the second column for each pure vowel sound, with mean duration only for the sounds *ar* (Group VII) and *er* (Group X). The fundamental and characteristic frequencies of each sound are shown in the 3 columns headed "Mean Fundamental," "Mean Low Characteristic" and "Mean High Characteristic Frequency" respectively. Each mean is taken from four records. The two columns headed "Scattered Low" and "Scattered High Frequencies" contain mean values of additional components, occurring in one or more records, in certain frequency ranges, the number of records in which such components are noted being shown in parentheses following the mean. The table illustrates and emphasizes many points which have been brought out in the preceding discussion, particularly the closeness with which the high frequency characteristics are defined in the vowels of the second or "doubly-resonant" class.

The table however gives no quantitative statement of the energy distribution among the different frequencies and it is necessary now to refer to the results of a harmonic analysis of these records which has been made and published[1] from which the diagram of Fig. 13 is taken. The machine method for analysing these wave-forms has been described by Mr. Sacia in detail elsewhere;[2] it suffices here to note merely the essentials in the treatment of the data.

For the dynamical study, the whole record from start to finish was taken as the unit for analysis, and the data obtained are therefore the average characteristics of the sounds throughout their duration. In the form of an endless belt each of these records was passed repeatedly through the analysing machine. A single record is of course

---

[1] "Dynamical Study of the Vowel Sounds." Bell System Technical Journal, III, No. 2, April, 1924.

[2] C. F. Sacia: "Photomechanical Wave Analyzer Applied to Inharmonic Analysis;" Jour. Opt. Soc. Am. and Rev. of Sci. Inst., 9, Oct., 1924, p. 487.

a non-periodic function, represented analytically by a Fourier Integral, not by a Fourier Series. The continued repetition of the record, however, builds up a periodic function consisting of a fundamental and a series of harmonics. The magnitudes of these components bear a simple relation to those of the infinitesimal components of corresponding frequencies in the Fourier Integral, and it is this series of relative amplitudes at different frequencies which is given by the mechanical analysis of the records.

It would be possible to present these results as the sound spectra of the vowels, showing their original acoustic pressure amplitudes[3] but this treatment has been modified for practical reasons to take into account the relative importance of the various pitches in hearing. Using the available data on the relative sensitivity of the ear at different frequencies[4] the pressure amplitude at each frequency has been multiplied by the corresponding ear sensitivity factor and the resulting curves are taken as the *effective* amplitude frequency relations which are most generally characteristic of these sounds.

The data from the four male records and from the four female records of each sound are separately averaged and the resulting curves are shown in the diagram (Fig. 13). This averaging process was somehwat laborious because the analyses of the separate records were made not with reference to predetermined frequency settings, but rather for those critical frequencies which best determined the shapes of the spectrum curves. The individual curves were therefore plotted on the musical pitch scale and the average ordinates were then read off for small intervals of pitch. These ordinates were then averaged for each group of four analyses. These average ordinates (after being corrected for the calibration of the recording apparatus) were then multiplied by the ear sensitivity factors for the corresponding frequencies. Thus the final spectrum diagram shows the relative importance of the amplitudes of all the components of each vowel for male and female speakers.

The amplitude units are entirely arbitrary; it is only the shapes,

---

[3] In Fig. 1, data have been given showing the actual distribution of energy in average speech. The tremendous concentration of energy in the lower frequencies is somewhat misleading unless account is also taken of the much reduced sensitivity of the ear in this region.

[4] See Bell System Tech. Journal, Vol. II, No. 4, October, 1923. The paper on Audition, by H. Fletcher, shows a graph of the "Threshold of Audibility" curve from which these data were obtained. The ear sensitivity factors used, of course, relate to the lower intensity levels; but it is thought that no essential inaccuracy is thereby introduced, as the position of the characteristic frequencies of a given vowel is subject to some variation with different speakers, and moderate variations in the height of these maxima in the energy spectra are not significant, except when taken from cycle to cycle in the case of an individual sound.

not the sizes of these curves which are significant. The order in which these curves are arranged is based upon the vowel triangle, and on Table IV. To return to the general discussion, we find that the fundamental voice frequencies do not have large effective amplitudes; it is interesting to note that these can be largely eliminated without impairing the distinctive quality of a vowel sound. The "scattered low frequencies" of the table (Sounds I to VII) exhibit appreciable amplitudes in the diagram. The "Scattered High Frequencies" of sounds I-VII previously noted exhibit small amplitude in the diagram. These are perhaps not essential to these speech sounds, but we should expect to find them in well trained singing voices. They are to a certain extent (particularly for the male voices), paralleled by the high-frequency regions of resonance for these sounds given in Paget's diagram, to which reference was made in Section I. Paget, it must be noted, is convinced that these high frequency regions of resonance are characteristic of the sounds of Groups I-VI.

The sound *a* (No. VI) is as it were the center of gravity of the vowel diagram and occupies the key position in the phonetics of most languages. The broad feature of the diagram is of course the progressive rise in frequency and gradual narrowing in range of the characteristic region of resonance, till the sound *a* is reached, succeeded by a splitting up into two regions of resonance which recede from one another as we follow the diagram downwards from *a* to the end. The exact location of sound X (er) is somewhat indeterminate, but it is evident that it belongs in the series of doubly resonant vowels. It is interesting to note that the distribution of the components of *ar* (refer either to Table IV or Fig. 13) is similar to the distributions given by Miller and by Paget for a form of the vowel *a* having "double" resonance; it is therefore as well located as any vowel in the series.

The characteristics of the *r* sound (whether considered as vowel or consonant) offer an interesting study, and in considering them we have an illustration of the practical value of records of the type shown. The problem of pronouncing a pure *r* sound is difficult; *r* is probably as variable in quality as any sound in the language, and it differs more than any other sound from one language to another. The precise location of its characteristic frequencies is thus a rather difficult matter. The records of *ar* and *er* disclose a noticeable tendency in speaking to make these sounds into diphthongs, the earlier portion of the record being nearly a pure *a* or (short) *e* while the latter portion of the record increasingly displays *r* characteristic. One speaker (MA) succeeded in making records for these two sounds which have nearly the same character throughout (Plates 49, 73), but for the other seven speakers, the "*r*" characteristics are best displayed toward the end of the record, though there is no sharp transition point. In the statistical study of these sounds the data were taken from the latter portions of the records; but in the mechanical analysis it was thought best to use the whole record. Now abstracting and condensing the data obtained in these two ways we have (ignoring fundamental tones) the following table of frequencies:

r (*ar* and *er*)

	From Table IV		From Fig. 13	
	Male	Female	Male	Female
Low	*570–630* 917 (ar)	*701–712* 1012 (ar)	*483–574* 861 (ar)	*512–542* 861–861
Middle	.....	.....	*1218–1448*	*1218–1448*
High	*1688–1965*	*2162–2188*	*1933-2896*	.... 1625 (er) *2435–2435*

These may be compared with Paget's results (from the second memoir, in which *r* is classified as a consonant sound) taking one of his general results from a mass of experimental data:

r (Paget: reference 9a, 9b p. 154)

"Throat or back resonance"	400-700 cycles
"Middle resonance"	1149-1824 cycles
"Front resonance"	1824-2169 cycles

(all varying with the associated vowel)

The *italicized* values in the first table above indicate correspondences with Paget's data, and we conclude that these roughly define the *r* sound, in terms of the steady-state theory.

Before taking leave of the vowel diagram, we should note not only the location of the resonant ranges but also their extent, and their relative separation from other resonant ranges in order to arrive at essential characteristics of the vowel sound. In other words, the individual vowel quality depends not only on a certain characteristic region of resonance but on the relative pitches in case there is more than one region of resonance. This effect is clearly shown to some degree in every group save one (VII:r) in Fig. 13. It will be noted that for the characteristic maxima of energy in the spectrum of a given sound, the peaks in the curve for female voices tend to occur at a higher frequency than the corresponding peaks in the curve for the male voices; but the musical interval between characteristic peaks for a given sound is about the same in the two cases. It is only in this way that we can account for what is a matter of universal experience in using the phonograph, namely that moderate variations from normal speed in recording and reproducing speech leave the vowel sounds still intelligible.

V.

Four Semi-Vowel Sounds[1]

Now consider the sounds *l*, *ng*, *n*, *m*, which pronounced with the vowels *oo*, *ee*, *a*, following them, are arranged in Groups XIV and XV. Following the plan previously used, note first the general characteristics of these 24 records, made by the two male speakers MA and MB. An outstanding feature of the records is the diphthong quality which is clear in all: the transition is quickly made from semi-vowel to the affixed vowel sound and except in two records (Plates Nos. 108 (*lee*) and 113 (*ngee*) a definite transition point can be fixed. Marking this point for all records we find an average duration of 0.16 second for the semi-vowel sound, of 0.21 second for the vowel sound, mean total duration being 0.37 second. Noting the fundamental frequency in two locations, namely at the start and just before the transition point, it is found that there is a progressive rise in pitch during the record of the semi-vowel sound; this effect is in agreement with the individual characteristics of these two speakers previously noted in the pure vowel records. But in addition it is noted that the average fundamental for these two speakers (see Table V below) is somewhat below that previously used by them in the vowel records. (Refer also to Table III). This slight lowering of fundamental pitch may possibly be a characteristic of the semi-vowel sounds; and this effect occurs, as we shall see later, to a pronounced degree in the consonant sounds.

The amplitudes of these semi-vowel sounds are on the whole smaller than the amplitudes of the affixed pure vowel sounds, but some of them are surprisingly large. The low frequency characteristic of *l* is (for these voices) principally a third harmonic of the fundamental. With *n* and *ng* (which are nearly indistinguishable) the second harmonic becomes increasingly important, and in the *m* records it is very large. The high frequency characteristics of all four sounds lie between 2400 and 2900, falling somewhat as we pass through a sequence from

---

[1] A preliminary report has been made on the properties of these sounds, and their relation to the general vowel diagram. (Phys. Rev. 23, 1924, p. 309.)

TABLE V

Speakers' Characteristics, Semi-Vowel Sounds

Sound	Duration in Seconds			Mean Fundamental (Semi-Vowel)	
	Semi-Vowel	Vowel	Total	At Start	Before Transition
l	.16	.20	.36	100	107
ng	.16	.20	.36	101	104
n	.16	.22	.38	98	107
m	.17	.20	.37	100	105
Mean	.16	.21	.37	100	106

l to *m*. We have here, then, a group of doubly resonant sounds whose characteristic frequencies, whose amplitudes, and general behavior are such that they must be definitely related to the standard vowel diagram.

The amplitude frequency relations as obtained from a mechanical harmonic analysis, and corrected for the variation in sensitivity of the ear are shown in Fig. 14. The process of mechanical harmonic analysis has been outlined in connection with the vowel records, and the procedure was the same here, except that only the semi-vowel portion of the records was taken as the unit for analysis. The record for analysis was cut at the end of the last cycle before the transition point, and two profile copies of the semi-vowel wave were joined together in an endless belt which was passed through the analyzing machine.

Aside from the close resemblance between the frequency spectra of the four sounds the noteworthy feature of Fig. 14 is in the similarity between the *l* spectrum and that for *ee* as previously given in line XIII of Fig. 13. The essential differences are a slight increase in the importance of the low frequency characteristics, and the slight shift of all the resonant regions toward lower frequency, in passing from *e* to *l*, and on through the sequence *ng*, *n*, *m*. We may thus regard the chart of Fig. 14 as a logical continuation of the generally accepted chart of Fig. 13 and place the four semi-vowel sounds definitely in an extended vowel diagram, following in regular order the sound long *e*.

Sir Richard Paget has made the interesting statement that "all the consonant sounds are as essentially musical as the vowels, i. e., they depend on variations of resonance in the vocal cavity, and should be capable of being imitated in the same way, if their characteristic resonances could be identified and reproduced in models." It is interesting to compare some observations made by him on *l*, *ng*, *n*, *m*, and reported in his second memoir. Working according to the method previously described (§I) Paget has constructed resonators which, under certain conditions, will produce transient forms of the four sounds we are discussing. Their tone constituents are identified by him as follows:

RESONANT FREQUENCIES, SEMI-VOWEL SOUNDS

(Paget: Reference 9b)

	"Throat"	"Middle" (Nasal)		"Upper" (Oral)
l	228–406[1]	683 (faint)	.....	1625–1932[1]
n	203–228	683	1217–1366	1448–2169[2]
ng	203–228	541–724	1217–1448	2298–2579
m	271	.....	1217–1448[2]	861–1722[2] 2434–2579 (faint)

[1] Varying and finally approximating a characteristic region of resonance of the associated vowel.

[2] Varying with the associated vowel.

Studying Paget's results in connection with those of Fig. 14, we note that the energy spectra clearly show the "throat" resonances for all four sounds in the neighborhood of 256 cycles. In the case of *n* the nasal resonance at 683 cycles (Paget) is one of the prominent tones centering around a frequency of 512 in the spectrum diagram. This resonance also appears prominently in the spectrum for *m* though Paget did not notice it. The higher middle resonances (1217-1448 cycles) which appear in Paget's table for the last three sounds appear also in the spectra for these three sounds according to Fig. 14. Allowing for the variation stated in notes (1) and (2) above, it appears that the upper (oral) resonances for the four sounds, as noted by Paget, are essentially the same as those that appear in all four spectra in the diagram in the range of 2048-2896 cycles.

With regard to Paget's observations on the transient character of these sounds (he classifies them as consonants) and on the variability of some of their components (Notes 1 and 2 of table above), depending on the associated vowel, there is room for some difference of opinion and the reader may form his own conclusions after a detailed inspection of the records shown. Taking the sound l for example, and studying first the three records *loo, lee, la* by M A and then the three corresponding records by M B it seems to the writer that such variations as are noted in characteristics are due not so much to change in the associated vowel as to the change in the speaker, and a similar conclusion will probably be reached for each of the other three semi-vowel sounds.

From the evidence in the records, it is difficult to subscribe entirely to a "transient" theory of these sounds, at least when they precede the standard vowel sounds. The evidence justifies the use which has been made of the steady-state idea, and the harmonic analyses leading to a determination of characteristic frequencies. But there is a possibility that the harmonic analysis does not tell the whole story. These two groups of records and the acoustic spectra based on them furnish outstanding examples of the niceties involved in speech and hearing in order to achieve the miracle of articulate speech. Without harmonic analysis, the most casual observer will note, for example, the similarity between the corresponding records of the *l* and *n* sounds, but more astonishing still is the resemblance between the *l* and *ee* sounds shown together in Plates Nos. 107 and 108. In this latter case (*l* and *ee*) practically the same high and low characteristic frequencies are involved, and it would seem that the distinction, which is sufficiently pronounced to the ear, must be based to some extent not only on the relative amplitudes of these frequencies present, but also on the behavior of these amplitudes during the fundamental cycle. It will be noted in practically all of the records of these semi-vowel sounds that the high frequency characteristic is a transient of more rapid decay than in the case of the pure vowel sounds; it is not of large amplitude except at the beginning of the cycle. On the face of the records this is the only explanation available for whatever distinctive quality these sounds, as a class, must possess.

VI

SIXTEEN CONSONANT SOUNDS

The last two groups, XVI and XVII contain, respectively, records of the "hard" and "soft" consonant sounds, each with the *a* sound affixed, and pronounced by the two male speakers. Here the classification is somewhat arbitrary; it is difficult if not impossible to arrange the sounds of these two groups in any such satisfactory series as has been determined for the semi-vowels of the two preceding groups. The sounds *dth* (that) and *th* (thin) for example have transitional characteristics that relate them to both groups; but they are placed at the end of Group XVI, to emphasize their relation to the pair *v/f* of the last group. With these reservations as to arrangement, consider the general characteristics of the consonant sounds of these two groups.

Examination first discloses a relatively easy separation of a given record into a consonant and a vowel portion and, as might be expected, a longer duration for the "voiced" consonants. In all the voiced consonants a sufficient portion of the record is reproduced to illustrate the voicing or fundamental of small amplitude in the early stages of the record; in the case of the unvoiced consonants of Group XVI this is not necessary. In the case of both the voiced and unvoiced consonants of Group XVII, longer records are shown, the high frequency component making this necessary, although the fundamental does not appear in the early stages of the unvoiced consonants of this group. The mean duration of the voiced consonants (*b*, *d*, *g*, *dth*) of Group XVI is 0.14 second; of the unvoiced consonants (*p*, *t*, *k*, *th*) 0.05 second. Aside from traces of the fundamental tone (and traces of its second and third harmonics) there is nothing of interest in the early stages of three of these four voiced consonants; in the case of *dth* there are traces of a high frequency (4200 and 2600 in the two records) in the early parts of the fundamental cycle. The voicing for all four sounds if uniformly of lower pitch than that used later in the records in speaking the vowel sound. Leaving the early stages, the record then proceeds to a transition point, lasting through from one to four cycles of the fundamental, and culminating in the appearance of the vowel sound. Before this transition point is reached, traces of high frequency appear in most cases, sometimes suggesting a single transient vibration. Aside from the lack of the fundamental vibration, there is a further distinguishing characteristic of the "unvoiced" sounds: a tendency of the first transition cycle of the fundamental to appear from 10 to 20 per cent shorter in duration than the mean of several following cycles. With both voiced and unvoiced sounds there is a tendency for a moderately low frequency (500 to 700 cycles) to appear during the transition; also a high frequency (of mean value 3225 cycles for the 16 records of this group) which latter may be due to the beginning of the *a* sound. Some of the individual characteristics of these records are given in Table VI.

The notable distinction between these sounds and the sounds of the next Group (XVII) rests on duration factors, and of even more importance, the pronounced high-frequency characteristics of the sounds of the last group. The mean duration of the voiced sounds in Group XVII is 0.21 second; that of the unvoiced sounds, 0.18 second. Two of the other characteristics are similar to those noted in the preceding group; first the voicing, where it occurs, is of abnormally low frequency, and second in the case of the unvoiced sounds, there is a marked shortness of the first fundamental cycle at the transition point. Except in the case of the sound *v* (Plates 145 and 146) the high frequencies are persistent and in many cases of large amplitude, both at the start and during the course of the consonant sound. These frequencies rise, as we go through this group, to values of 7000 and 8000 cycles in the case of the sounds *z* and *s*, shown in the last four records. For a full appreciation of these pronounced high frequency characteristics reference must be made to the records themselves, or the summary of characteristics, in Table VII. Here again, in distinguishing these sounds the remarkable performance of the ear is manifest, and the recording apparatus is used nearly to the limit of its utility.

We may best conclude this discussion of the consonant records by brief comments on some of the individual sounds, and a comparison where possible with data given for them in Paget's second memoir.

B/P.—(Plates 129-132). Both Paget (ref. 9b, p. 165) and Miller (ref. 3) have noted the essential impulsive quality of these sounds, and have produced them by sudden closing and opening of the mouth of a resonator. Paget considers *p* to be the more suddenly released, i. e. to have the steeper wave-front. From the records this is not evident; following the voicing period, the *b* would seem to be more suddenly produced, as judged by the growth in amplitude of the *a* sound following.

D/T.—(Plates 133-136). For both of these (see either Table VI or the records themselves) we note a high frequency characteristic of about 4000 cycles. Paget (9b, p. 168) observed "an upper resonance 5 to 8 semitones higher than that of the associated vowel, and a low resonance of about 362." We note in the records a low frequency of the order of 500 in the case of *d*. Paget notes a "greater amplitude in *t* due to higher air pressure" and the records show a greater amplitude for the high frequency in the case of *t*, except right at the transition point, where *d* shows the high frequency of large amplitude. No conclusion can be given as to relative steepness of wave-front, *d* vs. *t*, because in both cases we note for speaker MB (Records 134, 136) a steeper wave-front than for MA (Records 133, 135). The difference between *d* and *t* may depend entirely on the voicing and on the complicated phenomena at the transition point.

G/K.—(Plates 137-140). *k* shows the characteristic transients (1500, 4000; Table IV, notes 4 and 5) to much more pronounced degree than *g*. From the records it would seem that *g*, in addition to the voicing, disclosed a steeper wave-front, the *four* transitional cycles required for *k* (records 139-140) emphasizing this point. No other generalizations seem warranted, on account of the complicated series of events recorded. These sounds are treated at length by Paget (9b, p. 171-173) who observes considerable variation in their resonant ranges, depending on the associated vowel. It will be noted however, that in these four records particularly, consonant characteristics are persistent and of large amplitude before the vowel sound begins to appear.

DTH/TH.—(Plates 141-144). The high frequencies (2600, 3000, 3200) culminating at the transition point seem to be the key to these records. They are more persistent for *dth*, while *th* appears to show the steeper wave-front. Paget states (9b, p. 158) that "in δ [*dth*] the middle resonance [1149-1932, his figures] is overblown, - - - louder than the corresponding resonance in θ [*th*]." He gives also an "upper sibilant of 3444-5950," louder for *dth* than *th*, and "difficult to identify." It will be noted that in one record for *dth* (no. 141) there is during the voicing period a faint high frequency which has been set down in Table VI as 4000 cycles. This faint "sibilant" (which may always be audible though it fail to be recorded) establishes a certain kinship between these two sounds and those following (the fricative consonants) which are rich in sibiliant sounds.

V/F.—(Plates 145-148). *v* shows a pronounced voicing, and as previously noted, a less prominent high frequency component than its partner *f*, or any of the other fricative consonants. Comparing *v*/*f* with *dth*/*th* it seems from the records that the former pair are of higher frequency (particularly *f*) and that for *v*/*f* as a unit the high frequency characteristic is more pronounced; just the opposite conclusion to that reached by Paget (9b, p. 161-162). *f* may indeed differ more from *v* than *v* from *dth*, thus raising difficulties of classification both physically and phonetically, which cannot be resolved on the basis of the few records available. The exceedingly fine distinction between the sounds *v* and *dth* could be no more strikingly shown than it is in the records given, for both speakers.

J/CH.—(Plates 149-152). Some of the recorded phenomena of this pair suggest correspondences between them and the pair *g*/*k*; but the pair *j*/*ch* shows a higher frequency characteristic during the important mid-portion of its history. Of the pair, *ch* seems to show the steeper wave-front, that is, the more rapid transition to the vowel sound.

ZH/SH.—(Plates 153-156). With this pair we pass to the field of pure sibilants, in which there is no evidence of impulsive action or steepness of wave-front. The action seems to be that in the voiced sound, there is, in addition to the presence of the fundamental tone, a breaking up of the characteristic high frequency wave-train into discrete units corresponding to the fundamental tone, whereas in the unvoiced sound the high frequency characteristic is continuous, though irregular. Thus noting that the characteristic frequency is of 3000 to 4600 cycles the outstanding phenomena of *zh/sh* are well defined. In addition to frequencies of 2048-3249 noted by Paget (9b, p. 163) he gives a "pronounced middle resonance of 1625-2048." This latter observation of Paget's may correspond to the 1800-2000 frequency in the records of MB (Plates 154, 156) in the transition region, but this component does not seem to be prominent in the records.

Z/S.—(Plates 157-160). The general properties of these sounds can be inferred from the discussion of the preceding pair (*zh/sh*), adding only the fact that their principal characteristic is of much higher frequency. From Table VII we note a range of 4200-8000 cycles; Paget (9b, p. 162) gives "a characteristic upper resonance of 5790-6886." Paget also gives "a middle resonance of 1084-2298." The records do not show as low a range of characteristic frequencies unless it be the frequency range 2200-2800 (see Note 1, Table VII), within which fall certain vibrations occurring in the early parts of the fundamental cycles of the voiced sounds *zh* and *z*. The true *s* sound is, as Paget has stated, "a relatively complex hiss" and this is true of *sh* as well. And to complete the record, we must observe that *zh* and *z* are even more complex, if possible, and thus not inappropriate examples of the sounds of speech with which to conclude this survey.

To summarize, we have considered some of the more outstanding features of the wave forms of speech sounds which have been recorded. Many more detailed properties of these records deserve further study. The progressive change in wave form from cycle to cycle of the fundamental, particularly at the beginning of a sound, is undoubtedly an important factor in determining the character of speech sounds; it becomes most important, as we have seen, in the study of the more impulsive consonant sounds. There is material in these records for extended studies of this kind, which require a harmonic analyzer of a large number of components. We have not dealt with the question of the inherent power in speech sounds, another very characteristic property; these important data are accurately given in a paper by C. F. Sacia in this issue of the Journal. The relative power in consonant and vowel sounds can also be determined from those records in which vowels and consonants appear in combination, and it is hoped to carry this study further. Many other investigations of speech are now made possible on the basis of the accuracy of this set of records; in conclusion we may emphasize the fact that, for the present, the record is the important thing, and we believe that a set of faithful records opens a new prospect in the field of speech investigation.

TABLE VI

Group XVI—6 Stop Consonants; Transitional dth/th

Plate No.	Sound	Speaker	Consonant Characteristics					Transitional Characteristics				Vowel Fundamental	
				Near Start		Mid Portion to End							
			Duration	Voicing (Fundamental and Harmonics)	High Frequency	Voicing	High Frequency	Low Frequency	High Frequency (Note 6)	No. of Cycles	First Cycle Short	Near Start	Near End
129	*ba*	MA	.12	90,180	none	90,180	none	700	2700	1	. . .	100	115
130	*ba*	MB	.19	100,200	none	92,184	none	700	3100	1	yes	116	107
131	*pa*	MA	.02	unvoiced	none	unvoiced	2800 (Note 2)	1000	3600	1	yes	100	111
132	*pa*	MB	.04	(one 60 cycle vibration)	none	(One 60 cycle vibration)	3800	900	3600	1	yes	119	114
133	*da*	MA	.13	90,180	none	79,158	3800 (Note 3)	500	2800	3	yes	103	115
134	*da*	MB	.10	98,196	none	98,196	3600	600	3200	2	. . .	112	109
135	*ta*	MA	.07	unvoiced	none	(One 100 cycle vibration)	4300 (Note 3)	. . . .	3200	4	yes	104	112
136	*ta*	MB	.06	unvoiced	none	unvoiced	3600	900	3000	2	yes	120	113
137	*ga*	MA	.12	100,200,300	none	84,252	1600, 2800 (Note 4)	550	3000	3	. . .	101	111
138	*ga*	MB	.10	100,200,300	none	95,190	1400, 4000	600	3600	2	. . .	112	112
139	*ka*	MA	.07	unvoiced	none	unvoiced	1500, 4000 (Note 5)	1200	3800	4	yes	109	118
140	*ka*	MB	.08	unvoiced	none	unvoiced	1600, 4200	1300	4000	4	yes	125	116
141	*dtha*	MA	.20	83,166	4000 (Note 1)	95,189	4200 (Note 1)	600	3000	2	. . .	104	116
142	*dtha*	MB	.18	100,200	2600	100,200	2700	600	2600	4	. . .	109	107
143	*tha*	MA	.02	unvoiced	none	unvoiced	none	600	3200	1	yes	110	110
144	*tha*	MB	.02	unvoiced	none	(one 100 cycle vibration)	none	600	3200	1	yes	113	107

NOTE 1—A trace of these at beginning of the early fundamental cycles.
NOTE 2—One faint transient.
NOTE 3—Transients; longer for *ta* than for *da*.
NOTE 4—One transient.
NOTE 5—Irregular transients.
NOTE 6—Possibly due in some cases to the *a* sound.

TABLE VII

Group XVII—Fricative Consonants

Plate No.	Sound	Speaker	Consonant Characteristics					Transitional Characteristics				Vowel Fundamental	
			Duration	Near Start		Mid Portion to End							
				Voicing (Fundamental and Harmonics)	High Frequency	Voicing	High Frequency	Low Frequency	High Frequency (Note 3)	No. of Cycles	First Cycle Short	Near Start	Near End
145	*va*	MA	.20	97,195,390	3000	87,174	none	600	2700	3	. . .	101	116
146	*va*	MB	.25	112,224	3200 (trace)	100,200	none	600	3400	2	. . .	112	107
147	*fa*	MA	.15	unvoiced	3100	unvoiced	3500, 7000	500	2800	4	yes	112	121
148	*fa*	MB	.30	irregular	3200, 6400	unvoiced	3200, 6400	600	3600	3	yes	111	104
149	*ja*	MA	.22	81,243	3400	81,162	2600, 5200	450	2700	4	. . .	110	110
150	*ja*	MB	.14	trace	3300	90,179	2000, 4800	500	3100	4	. . .	115	111
151	*cha*	MA	.07	unvoiced	4800	unvoiced	2800, 4800	{500 1500	3000	2	yes	104	111
152	*cha*	MB	.08	unvoiced	3600	unvoiced	3600, 6400	{500 1600	trace	2	yes	119	115
153	*zha*	MA	.28	86,172,344	3000, 4000 (Note 1)	87	3000, 4000 (Note 1)	450	2900	4	. . .	100	111
154	*zha*	MB	.13	96	2600, 4200	99	3000, 4200	{500 2000	. . . .	4	. . .	114	111
155	*sha*	MA	.18	unvoiced	2800, 3600 (Note 2)	unvoiced	2800, 4600 (Note 2)	450	3200	3	yes	104	104
156	*sha*	MB	.17	unvoiced	2200, 5000	unvoiced	2600, 500	{500 1800	2800	3	yes	117	112
157	*za*	MA	.24	96,384	2800, 5600 (Note 1)	89,178	5200, 7000 (Note 1)	400	3100	4	. . .	98	108
158	*za*	MB	.22	100,300	2200, 4400	100,200	2800, 5600	550	2800	5	. . .	111	107
159	*sa*	MA	.27	unvoiced	5600, 8000	unvoiced	6000, 7800	500	2900	2	yes	114	114
160	*sa*	MB	.19	unvoiced	4000, 6400	unvoiced	4200, 6600	650	2900	2	yes	117	108

NOTE 1—Alternating; lower frequency in first part of fundamental cycle, higher frequency in latter part of cycle.
NOTE 2—Alternating, irregularly.
NOTE 3—Possibly due to the *a* sound.

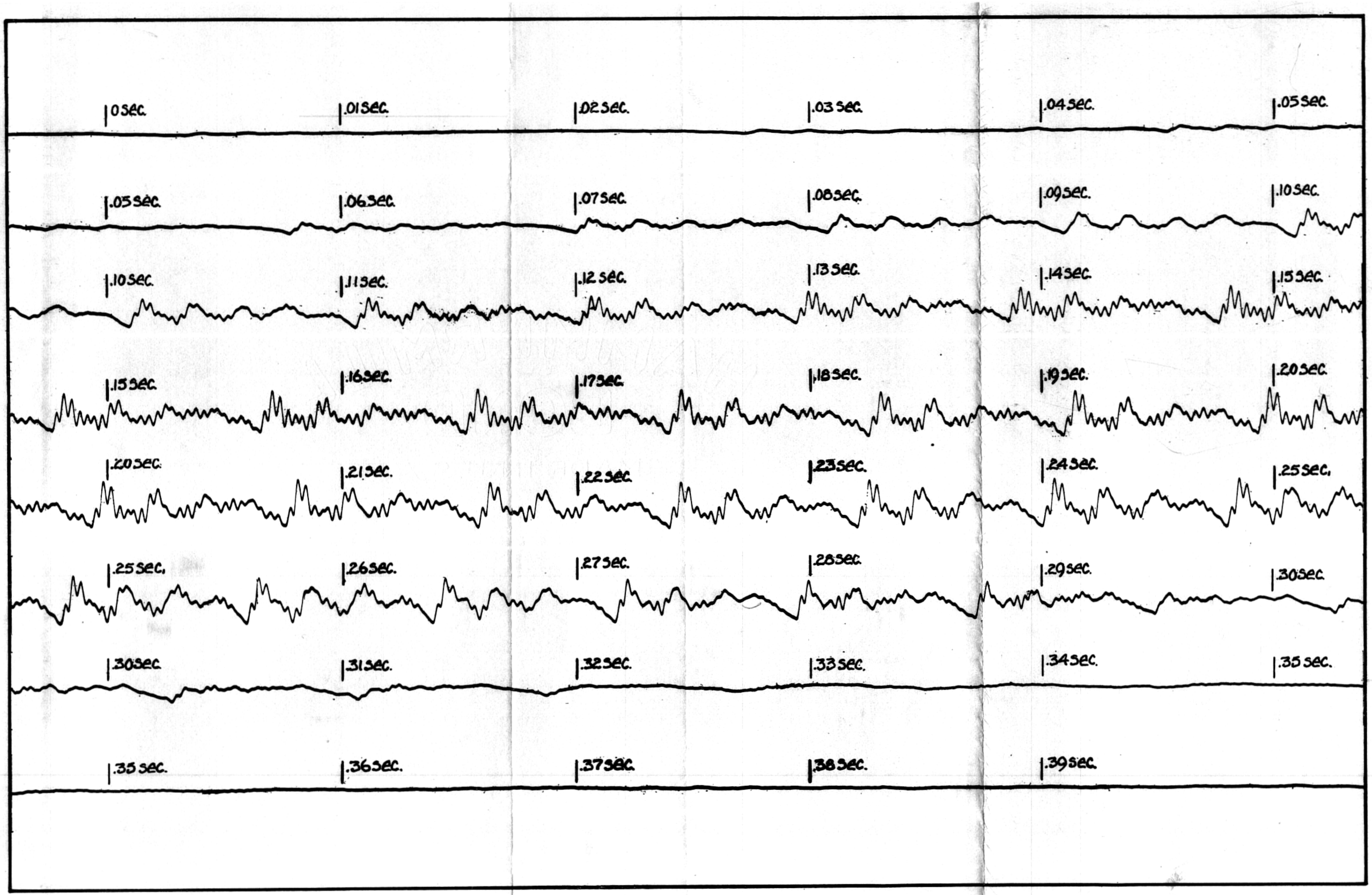


Plate No. 9—*u* as in *put*. Spoken by M.A.–Male, low-pitched (1)

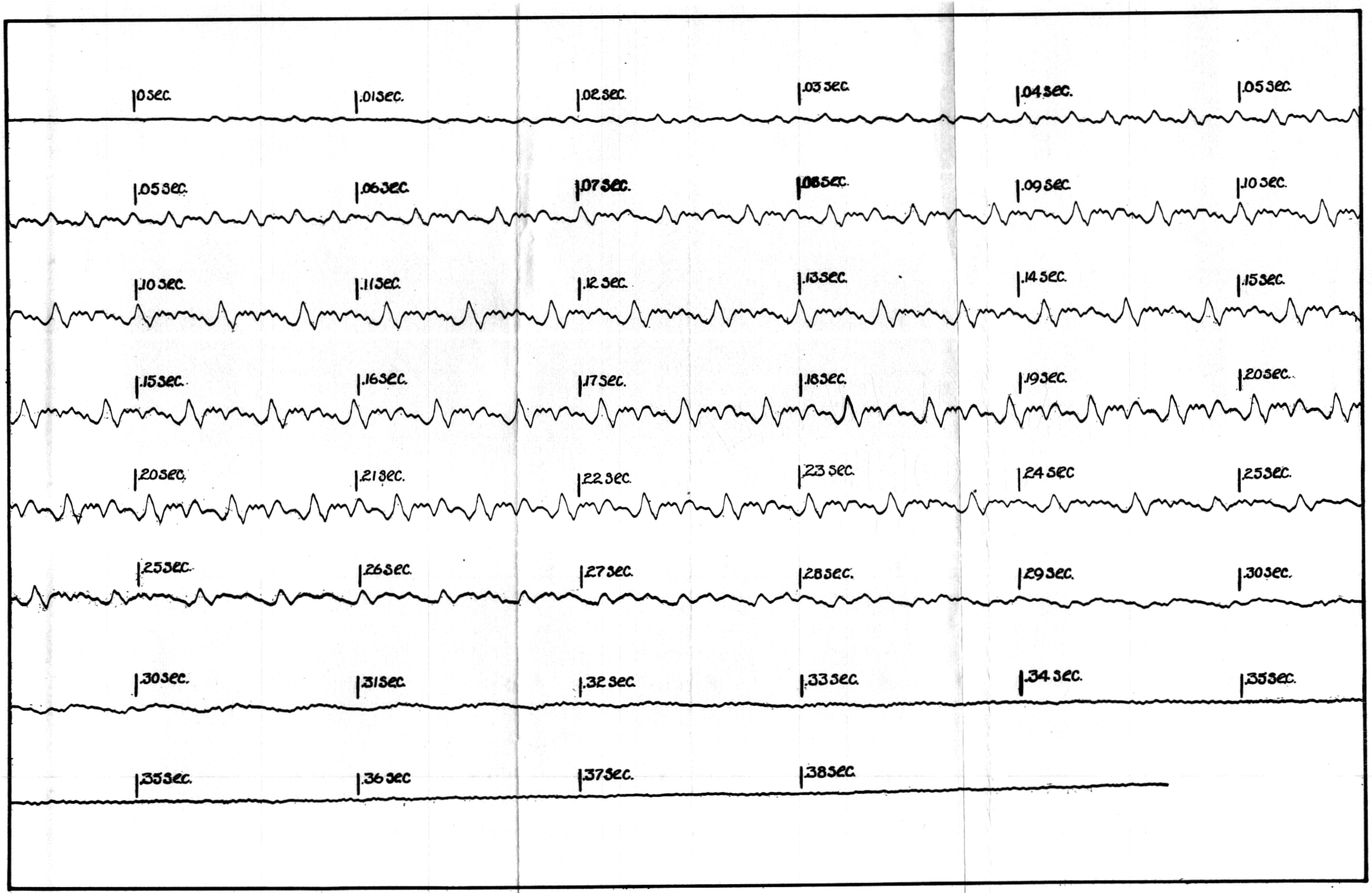


Plate No. 40—*o* as in *ton*. Spoken by F.D.–Female, high-pitched

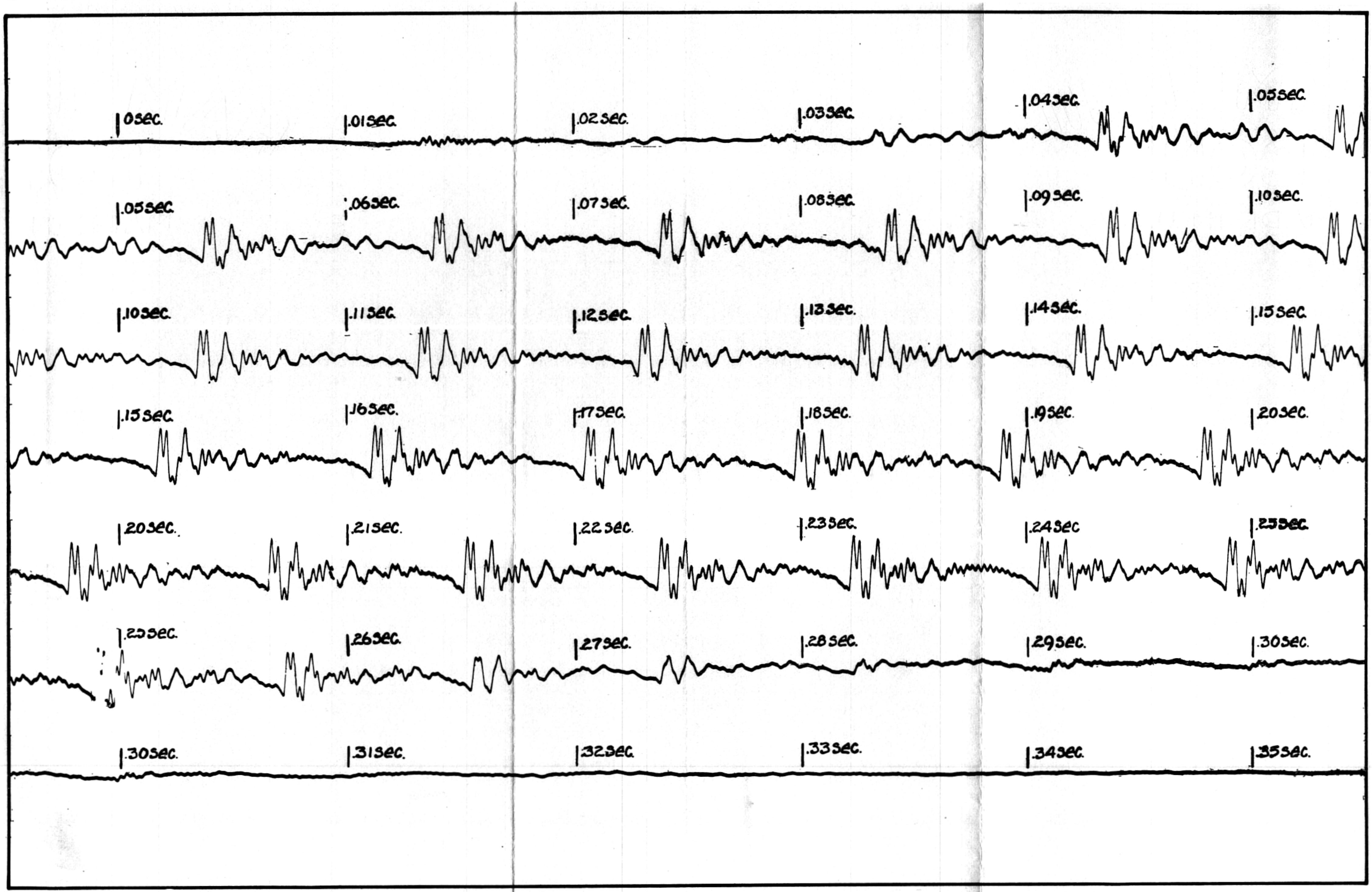


Plate No. 41—*a* as in *father*. Spoken by M.A.–Male, low-pitched

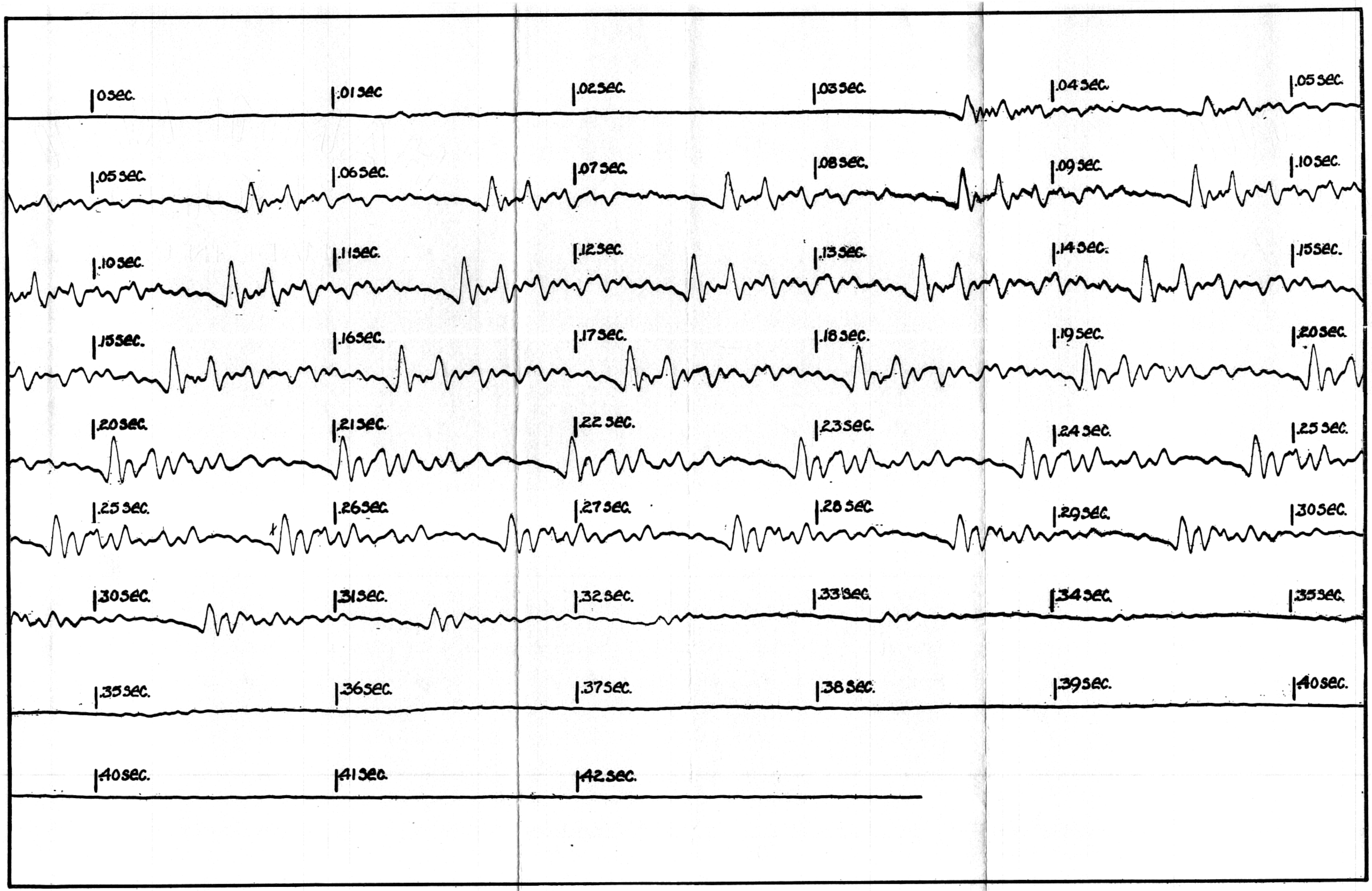


Plate No. 49—*ar* as in *part*. Spoken by M.A.–Male, low-pitched

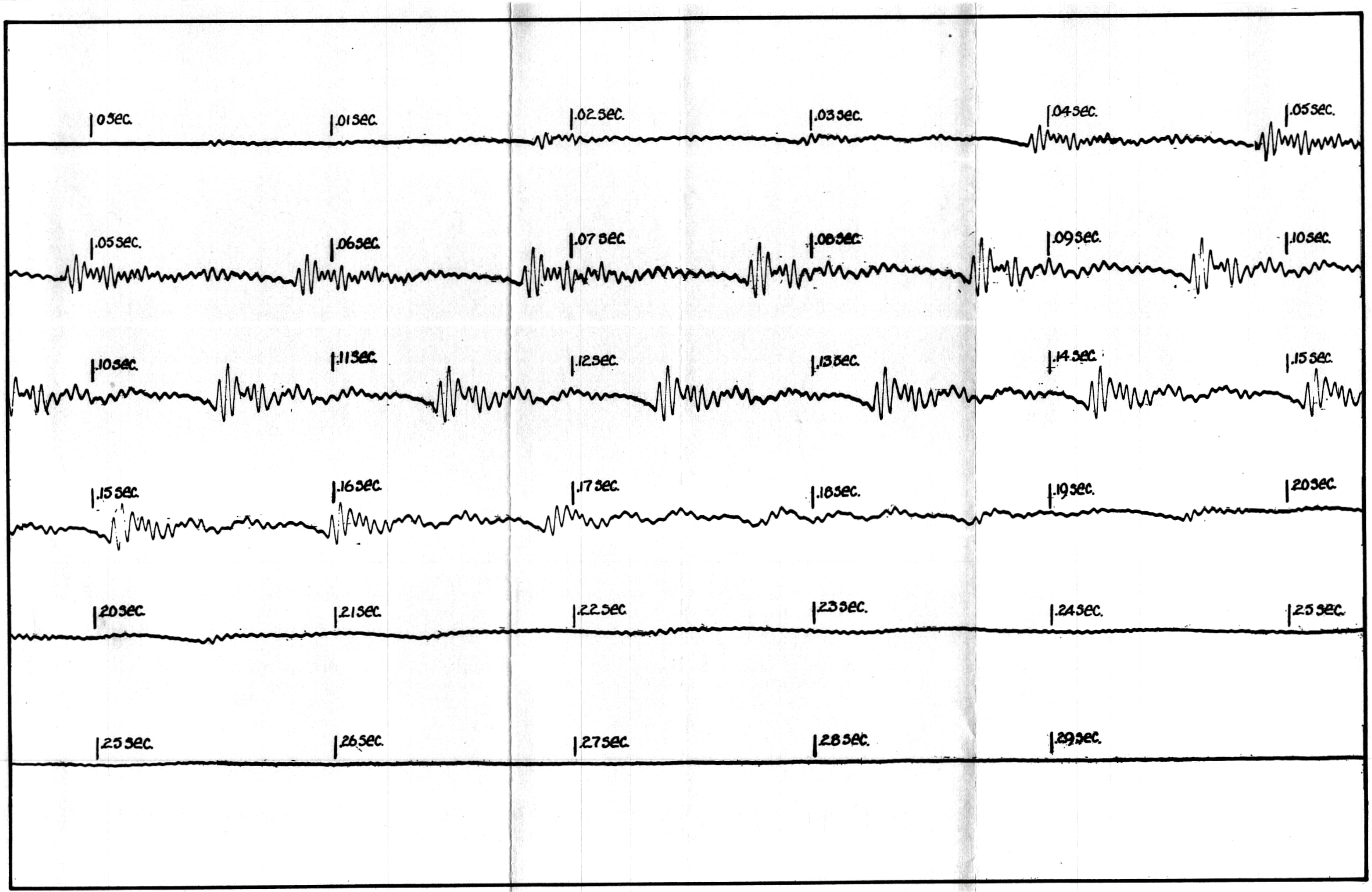


Plate No. 89—*i* as in *tip*. Spoken by M.A.–Male, low-pitched

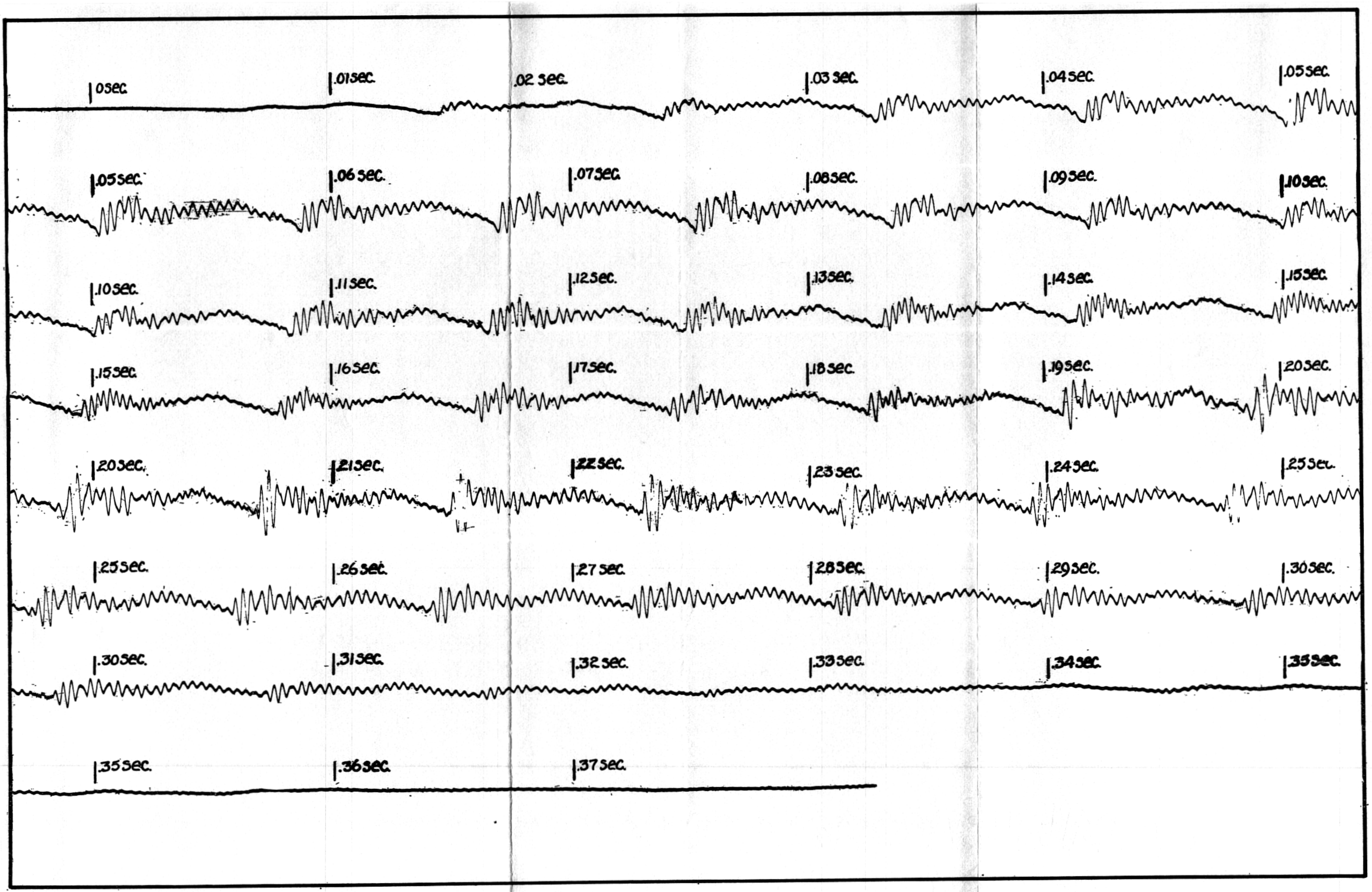


Plate No. 108—*lee*. Spoken by M.B.

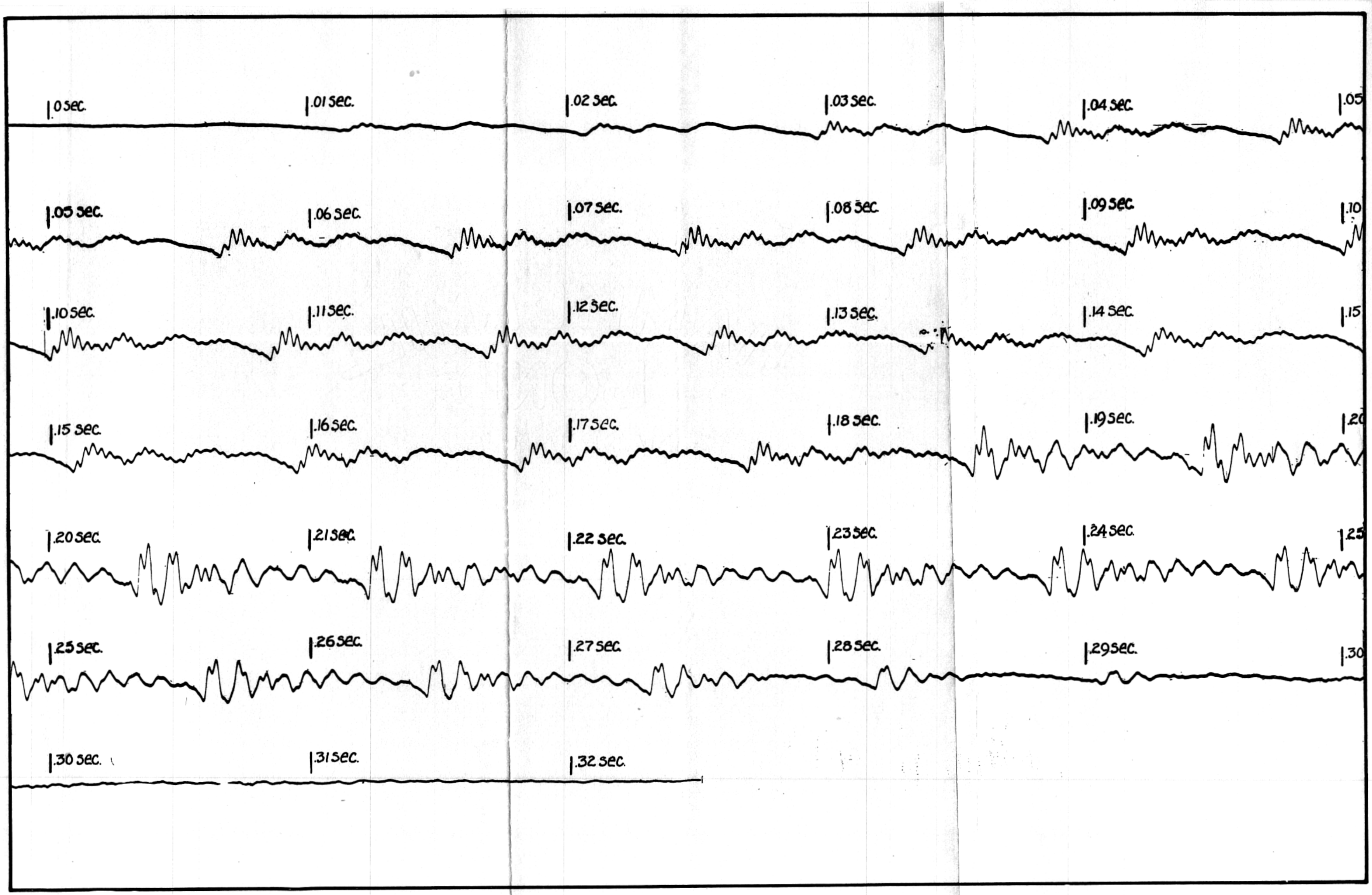


Plate No. 110—*la*. Spoken by M.B.

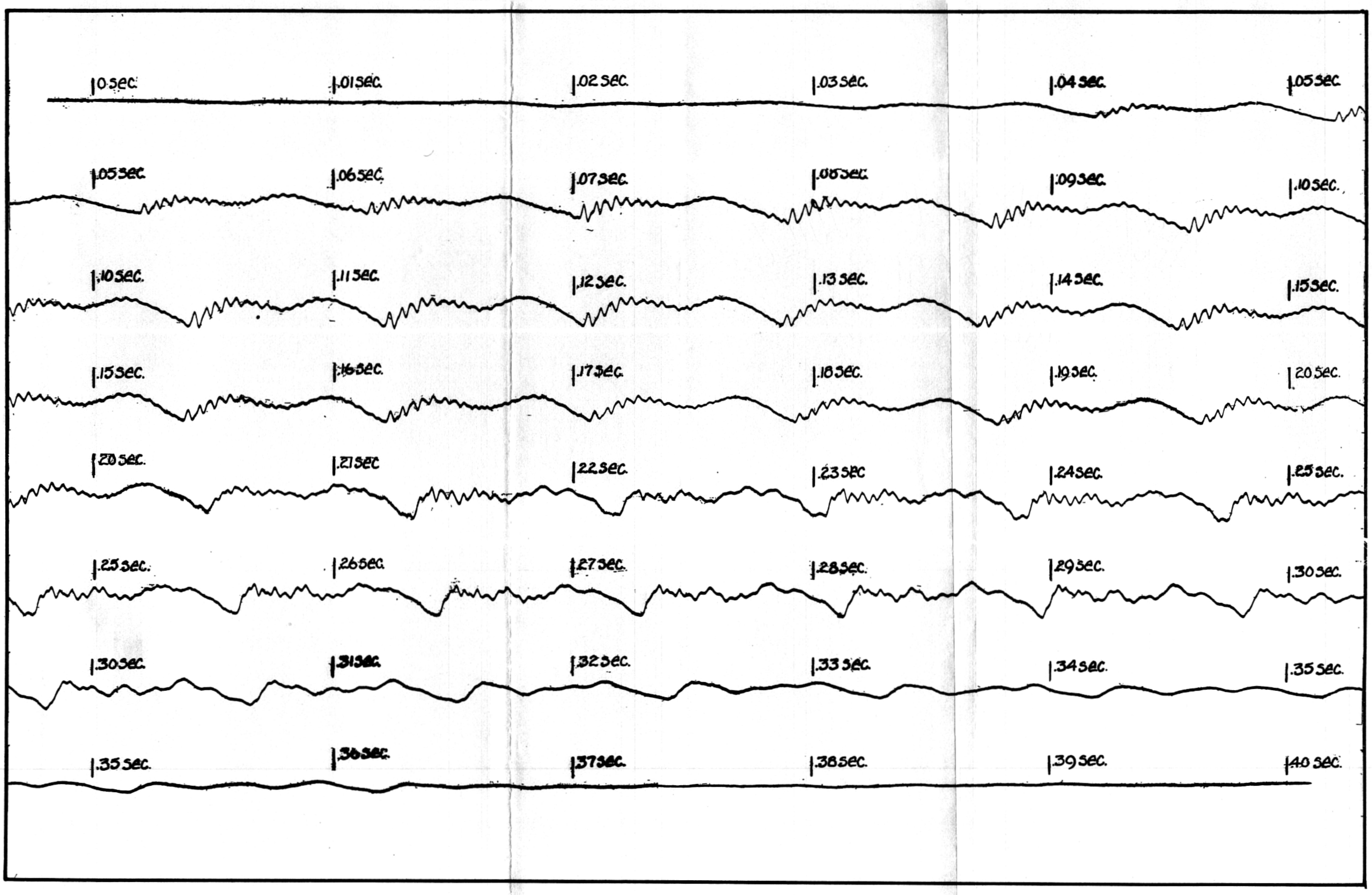


Plate No. 124—*moo*. Spoken by M.B.

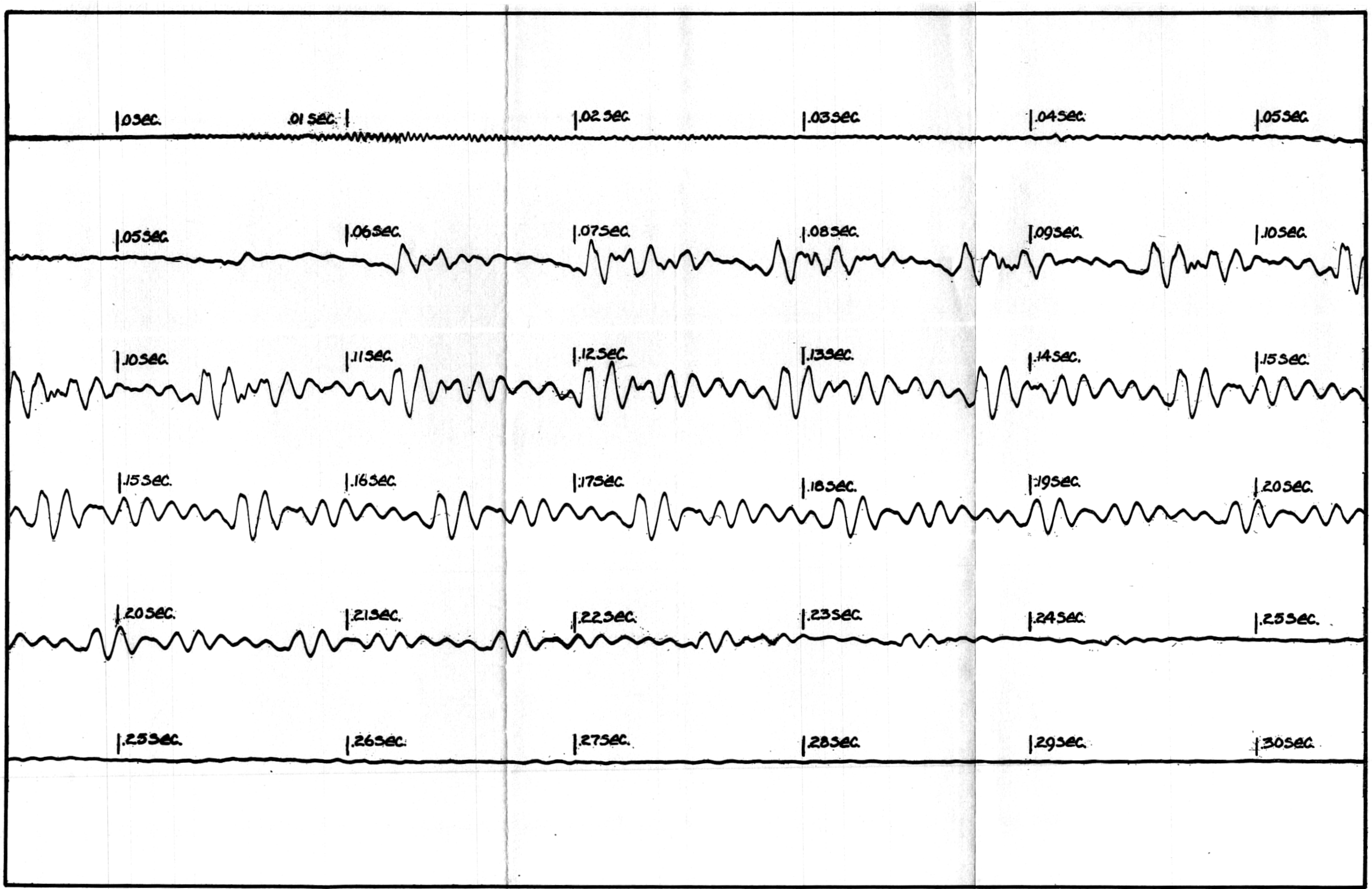


Plate No. 136—*ta*. Spoken by M.B.

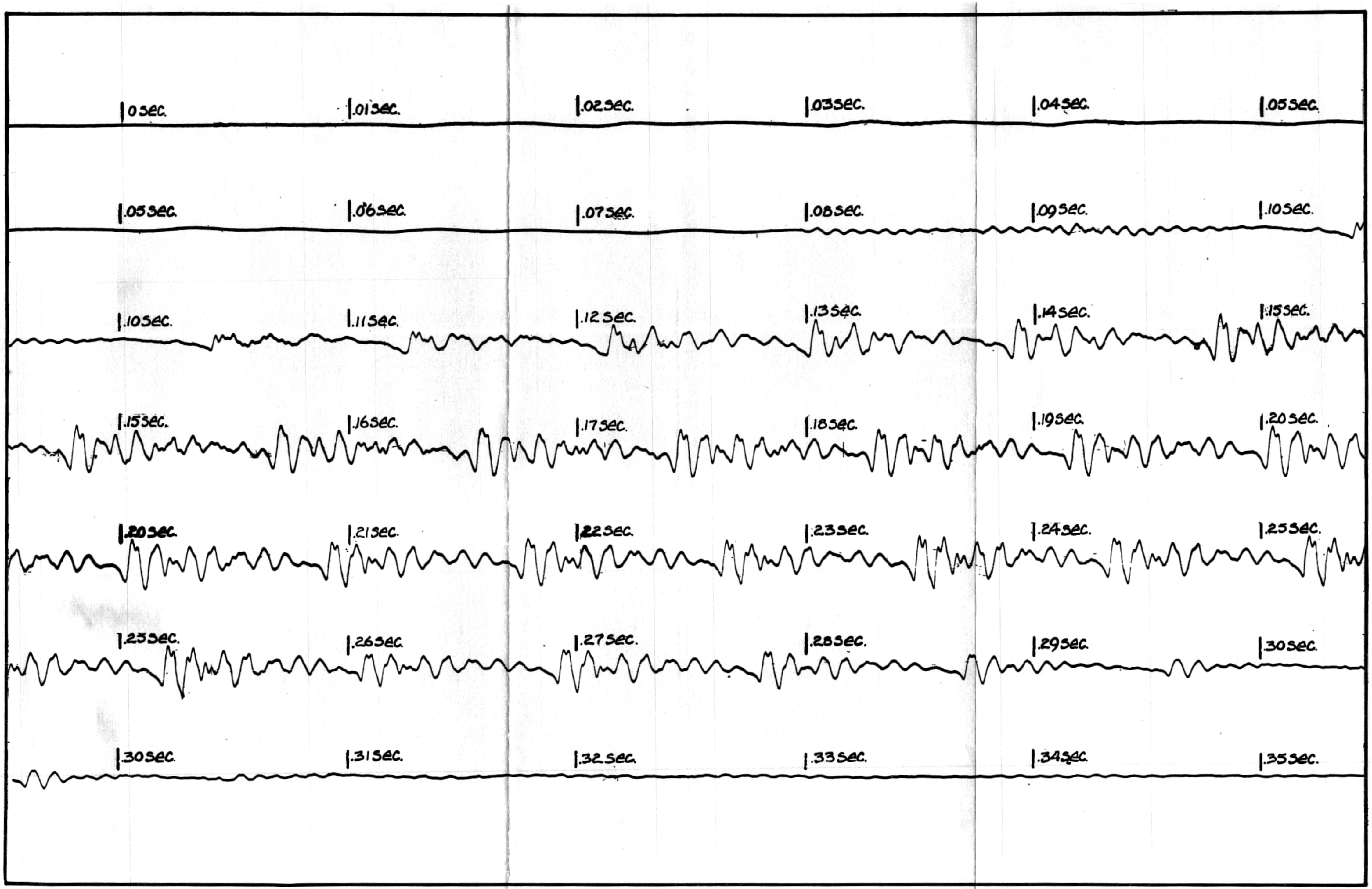


Plate No. 138—*ga*. Spoken by M.B.

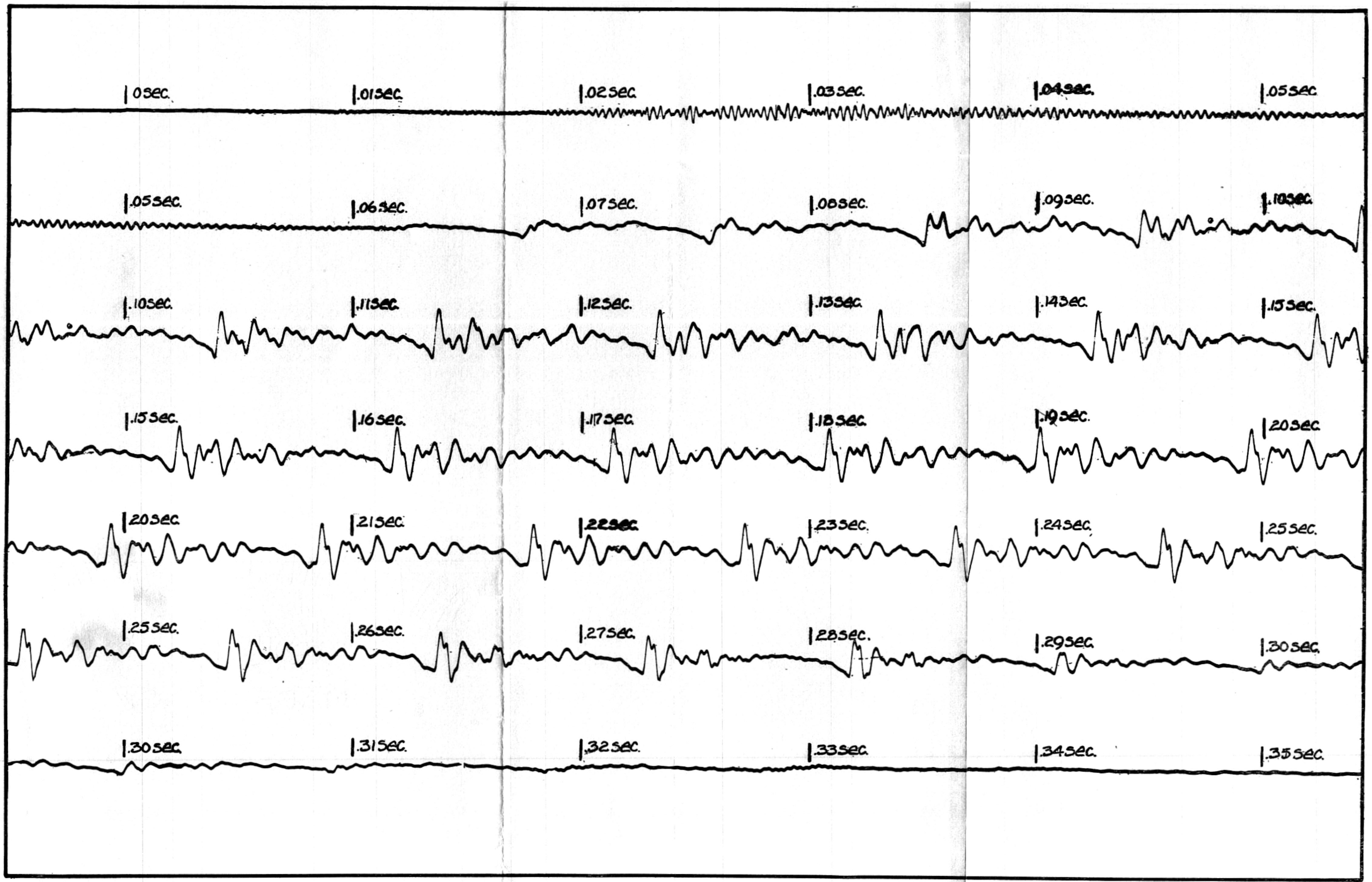


Plate No. 151—*cha.* Spoken by M.A.

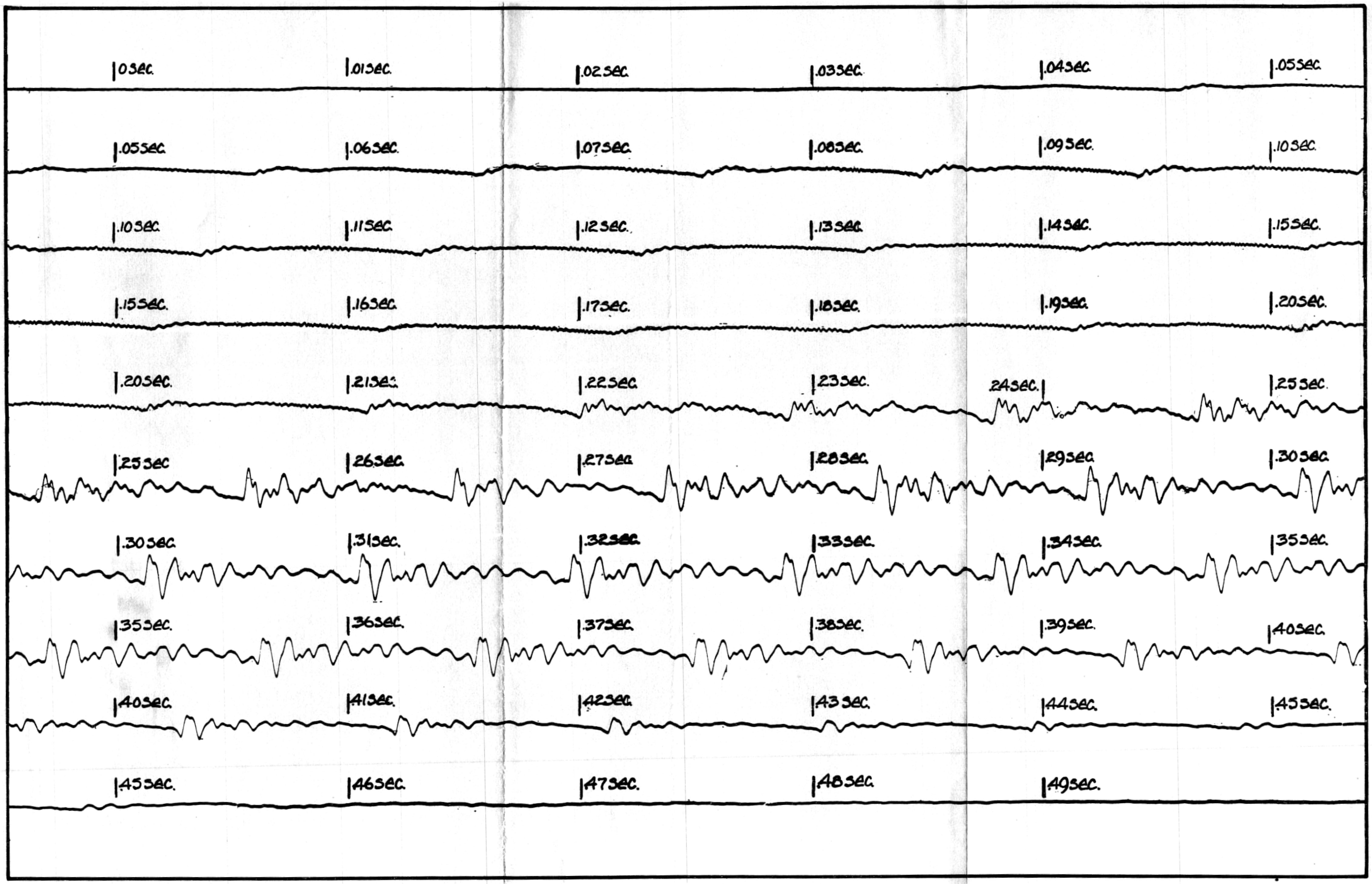


Plate No. 158—*Sa.* Spoken by M.B.

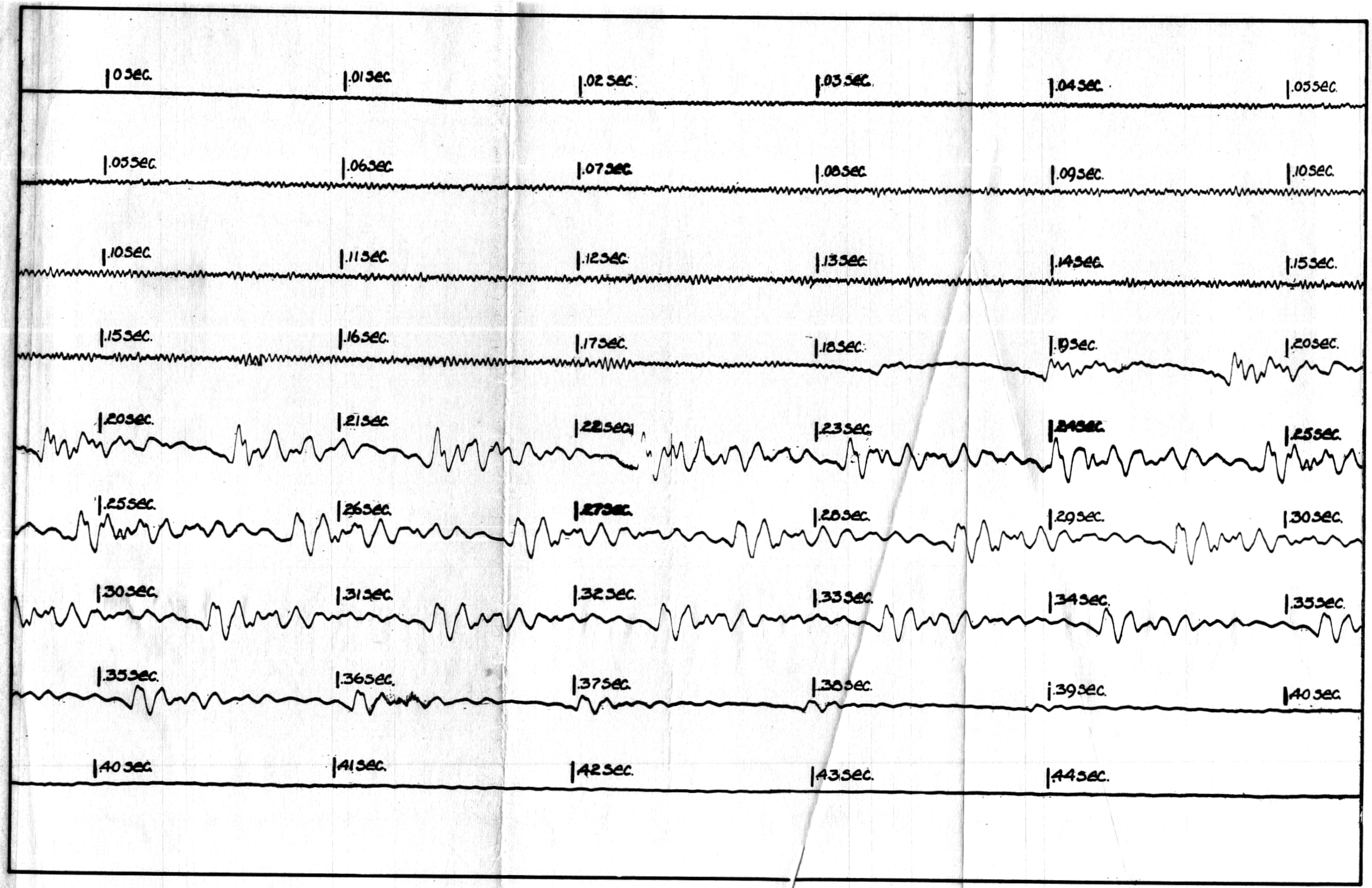


Plate No. 160—*Sa*. Spoken by M.B.