A VOICE PITCH INDICATOR FOR TRAINING DEAF SCHOLARS

THESIS

SUBMITTED TO THE

DEPARTMENT OF ELECTRICAL ENGINEERING

OF THE

UNIVERSITY OF CAPE TOWN

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

F. ANDERSON

1965

A VOICE PITCH INDICATOR FOR TRAINING DEAF SCHOLARS

SUMMARY

The instrument to be described, is based on experimental work documented by the author in 1960 and was developed specifically as a teaching aid to be used in the speech training of deaf scholars.

The only natural means by which a completely deaf child can compare his own speech with that of his teacher, is by observation of lip and facial movements and by feeling the vibrations of the vocal organs. Hence he is using the senses of sight and touch neither of which is capable of passing sufficient information to allow the child to develop good voice intelligibility.

Two properties of speech which contribute significantly to intelligibility are pitch and stress, and since these are relatively slowly varying quantities, the sense of sight can readily be trained to receive and process this information if presented to it in suitable form.

In this instrument, pitch or amplitude information is displayed as the ordinate of a graph, the abscissa of which is time. A continuous time-base is obtained by rotating a cathode-ray-tube with a long-persistence screen inside a stationary deflecting coil. The patterns thus formed, remain visible for a sufficient length of time for detailed interpretation by the sense of sight.

Pitch information is derived from measurements performed on the waveform of the speech signal, a process which unavoidably leads to errors. A system for detecting and eliminating these errors is described.

The application of the instrument, which has been used successfully over an extended period of time, is described briefly.

---oOo---

ACKNOWLEDGEMENTS

Documentation of this work was suggested by Professor R. Guelke of the Department of Electrical Engineering, University of Cape Town. Acting as supervisor, Professor Guelke's keen interest in the project served as a source of encouragement to the author and is gratefully acknowledged.

The work was done at the School for the Deaf, Worcester, by kind permission of the Board of Management and the Principal. The help of the teaching staff, with whom many fruitful discussions took place and who were responsible for the practical application and evaluation of the instrument, is highly appreciated as is the work of the typist, Miss J. Schreuder.

Finally, the author is indebted to the Council for Scientific and Industrial Research, Pretoria, for rendering assistance with the preparation of the diagrams and some of the photographs.

---oOo---

C O N T E N T S

ILLUSTRATIONS

---oOo---

A VOICE PITCH INDICATOR FOR TRAINING DEAF SCHOLARS.

CHAPTER 1.

SUMMARY OF ORIGINAL WORK.

The instrument to be described and its application to the speech training of deaf scholars, are based on preliminary work described by the writer in 1960 (1)* Throughout this treatise it will be assumed that the reader is familiar with this paper, which, for convenience, is briefly summarized below:

1. FUNCTION: The need was felt for an instrument which would portray in the form of visible patterns, the variations in the pitch of the voice, in order that a deaf child may have means for comparing his own vocal efforts with the voice of the teacher. Such a facility would enable him to experiment with his voice until he could produce the same patterns as the teacher. Being deprived of the normal channel by which speech is observed and controlled, namely hearing, and being forced to use the senses of sight and touch instead, a totally deaf person hardly ever learns to speak naturally. Imperfections in rhythm and stress often combine with unnatural pitch variations to lower the intelligibility of such a person's speech.

The instrument was designed to extract the pitch information from the complex sound signal falling on the microphone and then to display this as ordinate on a continuous time base. Periods of silence produced no display as did those portions of speech not produced by vibrations of the vocal cords. Only vocalized sound, produced by the action of the vocal cords, can be associated with the concept of pitch and hence only these sounds were displayed in order to keep the patterns produced by the instrument as simple as possible. This ensured that a deaf child could be taught to make efficient use of it with a minimum of preliminary training.

2. PITCH EXTRACTION: This was based on the "waveform approach" in which the beginning of the cycle, corresponding to the fundamental frequency component of the complex sound signal, was detected by a circuit which recognized two basic characteristics of the speech waveform, namely:

(a) if there are several peaks during the cycle, the first usually has the largest amplitude;

(b) the leading edge of this first peak normally has the greatest slope of all peaks.

The function of this circuit was to increase the amplitude of the first peak relative to subsequent peaks. By cascading several such shaping circuits, all but the initial peaks were suppressed. These were then further shaped into short pulses, each of which marked the beginning of a cycle at the fundamental frequency of the complex wave.

These pulses were then applied to a period-measuring device in which the time elapsed between pulses was measured. Measuring pitch period instead of the reciprocal, namely pitch frequency, resulted in an output which was for practical purposes an instantaneous analogue of pitch. Had the frequency been derived by the customary methods of time integration of pulse area, gross errors would have been introduced at each transition (e.g. silence to sound and vice versa) due to the time constant involved.

---

* A reprint of this paper is enclosed inside back cover.

3. DISPLAY OF PITCH ANALOGUE: The pitch analogue was displayed as a radial deflection of the spot of a cathode ray tube having a long persistence screen. This tube rotated around a horizontal axis, which was perpendicular to, and passed through the centre of the screen. Since the deflecting coil of the tube was stationary, the spot was stationary relative to the observer and when given a fixed deflection, left a circular trace on the long persistence screen as the tube rotated. By the time the tube had made one revolution, the previous trace had decayed below visibility and fresh information could be layed down by the electron beam. In the absence of voiced sound (i.e. during moments of silence and for the duration of unvoiced sound) the spot was extinguished so that for connected speech the resulting display consisted of an interrupted line whose radial position corresponded to the voice pitch. The length ofthe marks corresponded to the duration of the voiced parts of speech and these, taken in conjunction with the spaces between marks, constituted a vivid and easily interpreted pattern of speech rhythm.

A mask was placed over the screen covering all but a sector at the top centre of the tube. At the right hand edge of this window the C.R.T. spot excited the screen phosphor and, as the tube rotated, so the afterglow moved from right to left, allowing the patterns to remain visible for several seconds before disappearing from view behind the left-hand edge of the window. The patterns could therefore be studied briefly.

The radial position of the spot was made to be a logarithmic function of the pitch. Consequently equal (musical) intervals corresponded to equal steps in deflection. High pitched and low pitched voices uttering the same phrase, therefore produced similar patterns, the one high up on the screen, the other low down but directly comparable.

4. SUPPRESSION OF DISPLAY ON UNVOICED SOUND:

The criterion used to determine whether a sound was produced by the vocal cords or not, was that these two classes of sounds normally have widely different energy distributions in the frequency domain. Voiced sound generally shows a peak below 800 Hz, whilst unvoiced sound has an energy concentration above this frequency. Using low-pass and high-pass filters, the two general classes of sounds could be separated and the filter outputs could, after rectification and smoothing, be subtracted to yield a difference which was of one polarity for voiced sound and of the opposite polarity for unvoiced sound. It was applied to the C.R.T. to control the electron beam intensity. In addition a neon lamp was arranged to glow on unvoiced sound.

5. CIRCUIT DETAILS: The circuit diagram of the instrument was not given in detail, but was described with reference to a block diagram of the component parts. The operation of the period timing device was described in terms of its equivalent circuit, and details were given of the waveform shaping circuits. In a line drawing the mechanical details were shown and numerous oscillograms, showing typical waveshapes throughout the circuit, as well as typical pitch patterns as displayed on the revolving C.R.T., were presented.

6. DESIRABLE IMPROVEMENTS: After the experimental apparatus had been subjected to a series of practical tests, some shortcomings were disclosed and some improvements suggested themselves as follows:

(a) The patterns were not sufficiently bright and the phosphor decayed too rapidly for really satisfactory daylight viewing. It was thought that a tube should be sought which could retain a brighter image.

(b) If an image storage system could be developed which would retain the patterns at substantially constant brightness for a sufficient time, the possibility of arresting the motion of the time base would become attractive since this would allow the user to study the patterns in detail.

(c) If sufficient image brightness were available, it might become possible to apply intensity modulation to the C.R.T. spot in order to portray the other important property of speech, namely stress (amplitude).

(d) Errors occurred in the pitch extracting process at some of the transitions and also on sustained sounds in the case of a low male voice. Since pitch measurements were performed on a cycle to cycle basis and the display followed these instantaneous measurements without any smoothing or averaging devices, errors occurring even in only occasional cycles showed up in a very obtrusive manner and could hardly be ignored visually (octave jumping.)

7. <u>PRACTICAL USE AND RESULTS:</u> It became apparent soon after the instrument was put to practical use, that even the very simple patterns which appeared on the screen as the child spoke into the microphone, were not immediately meaningful to him. It was necessary to form simple associations between the sound vibrations which he could feel on his speech organs and the patterns produced by the instrument. The following steps were suggested as a training program:

(a) observing the duration of a sound;

(b) observing that only voiced sounds are portrayed on the screen, whilst unvoiced sounds cause the neon lamp to glow;

(c) observing rhythmic patterns without regard to pitch (marks and spaces in the display);

(d) observing that a stressed word in a phrase is often accompanied by a rise in pitch;

(e) when a child learns to associate the visible rise in pitch as portrayed with a corresponding tactile sensation in his speech organs, he is ready to start using the instrument in its intended role.

8. <u>CONCLUSIONS:</u> It was concluded that the experimental instrument had made possible the acquisition of experience leading to the development of a new design. It also served to demonstrate that it was a useful instrument in practice and that the children were interested in using it.

The instrument built according to this new design, will be described in the following chapters.

- - - oOo - - -

CHAPTER 2.

RECENT WORK IN THIS FIELD PERFORMED ABROAD.

A study of the literature reveals very little on systems of pitch extraction and portrayal between 1955 (2) and 1960, but since the latter date, several papers have been published on different aspects of these problems. Although some of these describe work not directly connected with the problem of designing an instrument for use with the deaf, it may be of interest to review briefly the work of others in this broad field.

1. GERMANY: In two papers (3) and (4) Kallenbach describes modifications to an instrument developed earlier by Grützmacher and Lottermoser (5). The system used to extract the pitch is based on the principle of enhancing the fundamental component by allowing intermodulation to take place between harmonics of the speech signal in a non-linear element. When the resultant signal is passed through a network, the transfer characteristic of which has a fall of 12 dB per octave above a certain "knee" frequency, it is found that the output signal consists mainly of the fundamental of the voice.* Harmonics are sufficiently suppressed to eliminate multiple zero crossings (as mentioned in Ref. (1), page 1066 under "Background") and thereafter it is a simple matter further to shape the waveform to produce one pulse per pitch period.

This pulse is used to trigger a saw-tooth oscillator, the free running frequency of which is below the lowest anticipated pitch. The saw-tooth wave is displayed on a cathode ray tube in the vertical direction and since the amplitude of the linear sweep is proportional to the period between triggering pulses, it represents the voice pitch.

No effort is made to provide a horizontal time scale by electronic means. Instead, the moving photographic film method is used. To accommodate different speakers, three different transfer characteristics are provided which have to be selected manually. Each of these allows operation over an octave.

COMMENTS: Since manual controls are deemed objectionable in an instrument to be used by technically unskilled teachers of the deaf, the system used to extract pitch information cannot be applied in the present case. The method used to obtain an instantaneous pitch display, is interesting since it is simple and the display, consisting as it does of a "wave envelope" type of pattern, the top of which is a straight line, the bottom representing the pitch, may not be less meaningful to a deaf child than the normal narrow line graph. The photographic method of display storage is, of course, not applicable in the present case, in which an instantly accessible display is essential.

2. ENGLAND: Plant - Mandy Voice Trainer.

This instrument, which has gained acceptance in some schools for the deaf in England, was developed by G. R. G. Plant of the non-profit organization known as "Speech Research for the Deaf Limited". The only description of this which could be traced, despite appeals to the author and the manufacturers for more details, is a non-technical paper (7) from which it was, however, possible to glean the following information:

(a) Throat microphones are used, presumably with the objects of improving the signal-to-noise ratio and of rendering the instrument insensitive to all speech sounds not asso-

---

* Essentially the same approach was used in a device described on page 57 of Reference (6).

ciated with vibrations of the vocal cords.

(b) Pitch information is obtained from the complex speech waveform by a process of filtering. A bank of overlapping one-octave filters is provided and of these one is selected manually which encompasses the pitch range of the person using the instrument. A stepped control selects filters with centre frequencies of 100, 125, 150, 175, 200, 250, 300, 350, 400 and 500 Hz and in addition there is a fine control providing a + 10% to - 10% adjustment in these ranges. Both these controls have to be correctly set for each voice to be displayed.

(c) The display of pitch and volume information takes place on a 5" cathode ray tube, apparantly not having a long persistence screen. When the voiced sound is of sufficient intensity, a mark appears, the position of which along the vertical axis is a function of pitch. If required, information about voice intensity is displayed as well "by arranging for the width of the light beam to expand laterally across the screen in accordance with the loudness of the speech sound being received." A selector switch is provided in order that pitch only, intensity only or both simultaneously, may be displayed.

<u>COMMENTS:</u> This instrument is , as far as can be ascertained, the only one of its kind specially designed for use in teaching the deaf, and as such merits some attention.

The use of a throat microphone may have advantages over the more usual type from a technical point of view, but its proper operation calls for strapping to the user's throat which, in practice, may be unacceptable, since it is frequently necessary to pass the microphone around to a number of pupils with a minimum of delay.

The necessity manually to select the proper octave range would almost certainly lead to difficulties in the hands of non-technical staff. Also, comparison between teacher's and pupils's voices would be difficult when a different range is required by these speakers, as will often be the case.

Experience shows that a display which does not provide a time dimension is of less value than one which retains the information of the past few seconds of time. In fact, the absence of a "memory" in the display was one of the main objections raised against the Coyne Pitch Indicator by its users.*

The facility of displaying pitch information and voice intensity information separately or simultaneously, is a valuable feature.

3. PITCH EXTRACTION BY COMPUTOR TECHNIQUES:

A. COMPUTATIONS PERFORMED DIRECTLY ON THE SPEECH WAVEFORM:

In a recent paper (8) Gold mentions that pitch extraction may be categorized as a problem in pattern recognition, other examples of the latter being the automatic recognition of handwritten letters, mechanical speech recognition and Morse decoding machines. A property common to these problems is that all can readily be solved by the human senses. These,in conjunction with the brain, form a sensing and computing system not easily duplicated outside the human frame.

Gold goes on to say that past efforts at pitch extraction <u>were aimed at singling</u> out one operation which would make pitch

* See Reference 1 page 1066.

extraction as simple as possible. These efforts were directed along two main avenues of approach:

(a) detecting the regularity of the waveshape in large sections of speech;

(b) detecting the "peakedness" in the waveshape of the sound produced by the puffs of air coming from the lungs via the vocal cords.

Circuits designed to recognize the first of these properties use filters to cut out unwanted harmonics and operate successfully over one octave and on sustained sounds, but fail on transitions. Since speech is a continual succession of transitions (abrupt changes in intensity, pitch or harmonic structure) errors occur too frequently to be ignored. Usually the voice pitch, even for a single speaker, has excursions of more than one octave which can lead to additional difficulties in the case of any system using filters.

Circuits based on the "waveform approach" detect peaks in the waveform and shape these to increase their amplitude relative to minor peaks. By repeating this process several times, a pulse train of pitch period markers which is reasonable devoid of spurious pulses is produced. Errors can occur when the waveshape is insufficiently peaked (according to Gold) at a time when the wave is perfectly regular and a pitch extractor based on filtering out harmonics would have performed well. Thus one type of pitch extractor may fail where the other succeeds and vice versa.

The computor system of pitch extraction described by Gold makes use of both the above properties of a speech wave to obtain a closer approximation to the pitch as detected by the human senses than is possible when using only one or the other. Gold successfully tested his theories using an available computor suitably programmed and did not propose constructing a selfcontained instrument on these lines.

B. COMPUTATIONS PERFORMED ON FOURIER ANALYSIS DATA:

In two recent papers (9) and (10) approaches to the problem of pitch extraction are described which are similar in some respects and which will be summarized collectively. The processes involved are analogous to those which probably take place in the human hearing system when judging voice pitch.

The basis of both systems is to perform first a normal Fourier-analysis of the incoming speech signal yielding a complete amplitude vs. frequency spectrum many times each second. Several different methods of achieving such a rapid analysis in real time have been described (11), (12), (13), (14), and another system is described as part of (10). All these methods are able to provide a frequency resolution of the order of 60 Hz and produce a spectrum analysis covering the range from about 60 to 5000 Hz within approximately 20 milliseconds. The fine resolution allows individual harmonics of even a low male voice to be resolved and the rapid analysing time allows the output to follow rapid changes in the spectrum such as those encountered in speech.

Up to this point the method used by Harris and Weiss (9) is in principle similar to that used by Noll (10), but the manner of finding the pitch from these data differs and will now be outlined separately:

(a) Harris and Weiss:

The spectrum is fed to a computor which is programmed to measure frequency differences between consecutive harmonics. This operation is performed at selected frequencies throughout the spectrum (usually at the positions of the formants) and the computor decides which measurements are most reliable, takes an average of these and prints out the result, the pitch being equal to the average frequency difference between successive harmonics.

(b) Noll:

It was pointed out as long ago as 1950 by Huggins (15) that the complex speech wave could be looked upon as the product of two functions: the excitation function (waveform produced by the vocal cords) and the system transfer function (the influence of the vocal tract upon the sound produced by the vocal cords.) Noll therefore concluded that, if a Fourier analysis of the speech wave is performed and the resulting spectrum is presented as the logarithm of amplitude vs. the frequency, then the product formed by the speech organs of these two functions would be displayed as a linear addition of their logarithms.

In this logarithmic spectrum, the harmonics of the vocal cord waveform (a buzz) are therefore superimposed as ripples of constant, short "wavelength" on top of a long "wavelength" envelope representing the system transfer function. Here the term "wavelength" should be understood to mean the distance between successive peaks of the curve as measured along the frequency axis (abscissa) of the conventional frequency spectrum

If this analogy is carried one step further and the abscissa is looked upon as having the dimensions of time and not those of frequency, then the original curve becomes simply a log-amplitude vs. time function. This may therefore be subjected to a second Fourier analysis in the usual manner to produce a spectrum of "frequency" components. Since the ripples of constant "wavelength" are clearly defined and regular throughout the original spectrum, they form a very large amplitude component in this second spectrum. This component represents the voice pitch since it is a measure of the frequency difference between successive harmonics in the frequency spectrum.

It follows immediately that on unvoiced sound (i.e. one in which the buzz of the vocal cords is not produced) these regular recurring harmonics will not be present and hence there will be no large amplitude component in the second spectrum. Hence voiced - unvoiced discrimination is achieved automatically in this system.

Furthermore, both these systems possess a degree of noise immunity, since noise usually has a random distribution over the frequency spectrum and is therefore easily separated from the single line component representing the pitch frequency. Amplitude or phase distortion of the speech wave does not destroy its periodicity and removal of low frequency components from the speech does not affect the higher frequency components of the vocal cord buzz. Consequently none of these common imperfections of sound reproduction have an important effect on the accuracy of pitch extraction.

COMMENTS ON 3. B: Similarities between these two approaches and the human sense of hearing are immediately apparent. In the ear the incoming sound wave forms a pattern on the hair cells of the basilar membrane. This constitutes, in effect, a Fourier analysis of the sound which is conveyed to the brain by the auditory nerve. The brain processes this information and possibly judges pitch in a way which is analogous to those described above, since it is also relatively immune to frequency, amplitude and phase distortion, as well as to interfering noise and to the removal of low frequency speech components.

As such, these systems yield pitch information with an accuracy of 1 or 2 Hz, and hardly ever produce gross errors at transitions or when dealing with some of the more difficult sounds of a male voice - a serious shortcoming of most other systems of pitch extraction.

Although undoubtedly the very ultimate in the field of pitch extraction, these systems call for a degree of sophistication which at present is difficult to accomplish in a unit which has to be self-contained, reasonably priced and capable of giving satisfactory performance when used by a technically unskilled person.

For the present application, a much simpler approach is therefore both justified and necessary, if the instrument is to be reproduced in any quantity, if it is to be useful as a teaching aid and if its price is to be attractive to prospective users.

--- oOo ---

CHAPTER 3.

STATEMENT OF THE PROBLEM.

As originally envisaged, the new instrument had to conform to the following specification:

(a) From the sound produced by the voice, the quantity known as the pitch should be extracted. Since pitch, as interpreted by the human hearing mechanism, is a function not of waveshape but of the fundamental frequency component of the complex wave, the pitch analogue extracted by the instrument should also be independent of waveshape.

(b) The pitch analogue should be displayed in visible form on a continuous time base which allows the information portrayed during the foregoing 5 seconds of time to be observed at a glance. The display should, as far as the user is concerned, present the visible information simultaneously with the sound. This implies that a time delay of perhaps one-tenth second is permissible. The pitch scale should be logarithmic.

(c) These stored patterns should remain visible for a sufficient length of time to allow the user of the instrument to examine them at leisure if he so desires. To this end it should be possible to arrest the motion of the time base.

(d) The display should be well visible when sufficient ambient light is present for comfortable lipreading, since with less illumination the deaf child who uses the instrument will be deprived of this, his normal means of communication with the teacher.

(e) Since pitch is only associated with vocalized speech (such as the vowels and nasal consonants) during the production of which the vocal cords are vibrating, only these sounds should activate the display. Unvoiced sounds such as sybilants, stop consonants etc., and silent periods, such as occur between some words, before plosives etc., should form blanks in the pattern.

(f) The visual pitch pattern should satisfy the eye as being a reasonably accurate replica of the pitch variations as interpreted aurally. Obviously the brain is able to disregard a great deal of irrelevant information in assessing the pitch, such as the harmonic structure of the sound, as well as noise in its various forms including reverberation (which is interference by a time delayed and frequency distorted replica of the desired signal). This amounts to stating that the hearing mechanism can operate successfully under conditions of poorer signal-to-noise ratio than those with which most practical instruments can cope. A favourable acoustic environment should therefore be provided and the instrument should be designed with a view to handling a large dynamic range without requiring manual adjustment (possibly by incorporating automatic gain controlled amplification in the microphone preamplifier.)

(g) The "unvoiced sound indicator" of the experimental model should be retained since it provides additional interest for the child. An attempt should be made to improve its performance on some of the more difficult unvoiced sounds, such as the soft plosives (e.g. poor).

(h) Tests conducted with the experimental model showed that some male voices went below 100 Hz and some children produced sounds at a pitch of more than 600 Hz. Investigation showed, however, that the three octaves from 90 to 720 Hz would suffice in the majority of cases. The new design should therefore be based on this range.

(i) The instrument should be reasonably transportable, preferably consisting of one unit only. It should have a minimum of controls calling for skilled manipulation, since it will be used by technically unskilled personnel.

--- oOo ---

CHAPTER 4.

DESIGN OF NEW INSTRUMENT.

It was recognized at the outset that many of the design details would have to be finalized empirically rather than by a rigorously scientific approach. Although the first experimental model of the pitch indicator was available to check certain new ideas, the new instrument could not be built as a complete unit from the drawing board without considerable further development. It goes without saying, therefore, that in its present form the instrument does not reveal some of the lesser successful ideas which were investigated and discarded. It is felt, however, that it may be wise also to document some of these, not only for the sake of completeness, but also since these appeared at first to be logical avenues of approach, but were later proved otherwise by practical experience.

PRELIMINARY DESIGN: This is discussed with reference to the block diagram of Fig. 1.

Block 1: Microphone input amplifier, preset volume control and phase splitter. The push-pull output of this stage drives the next.

Block 2: Variable-gain stage using two valves, having dual control grids in order to enable the functions of amplification and gain control to be separated. These valves operate in push-pull, as far as the signal path is concerned, and in parallel for the gain control voltage. Hence the control signal disappears from the output of this stage which is transformer coupled to the next.

Block 3: The A.G.C. rectifier and D.C. amplifier functions are combined in one valve operating under Class C conditions. A sharp cut-off pentode is used here. The voltage by which the bias exceeds cut-off bias approximates the A.G.C. delay voltage, as this valve has a very short grid base. Since the control voltage should be negative-going and near ground potential and since it appears across the anode load resistor of this stage, the valve is operated with a negative H.T. supply connected to its cathode. After smoothing, this amplified A.G.C. voltage is applied to the variable-gain stage and also to the next which is the signal gate.

Block 4: The function of this signal gate is to prevent signals of insufficient amplitude (which therefore have a poor signal-to-noise ratio) entering subsequent stages, since their presence may lead to false indications. In principle, therefore, either correct operation is obtained, or no operation at all.

Block 5: This stage merely serves to provide a low drive impedance for the subsequent stages and to raise the signal to the required level. It feeds three filters which separate frequency components carrying the pitch information from those representing the main energy content of the unvoiced speech sounds (Blocks 6, 15 and 20).

Block 6: This is a 90 to 720 Hz bandpass filter of the active RC type with 24 dB per octave slope beyond the cutoff frequencies. Its function is to remove from the complex speech wave as many of the components which are irrelevant to the determination of pitch as possible. It was experimentally determined that such filtering had no detrimental effect on the waveshape as far as those properties are concerned on which pitch extraction is based. The presence of strong components in this frequency range is used furthermore as an indication that a voiced sound is being dealt with - information which is further processed in Blocks 17 and 18.

Block 7: The three shaper stages are, in principle, identical and are connected in cascade so the shaping process takes place in three steps. The principle of operation is the same as in the first experimental model (see Ref. 1, page 1067), but an additional property of the speech wave was exploited, namely that the area under the initial peak of the pitch cycle is usually larger than the area under subsequent peaks. Hence a small amount of integration could be employed to improve the shaping process. The final shaper stage produces an output which normally contains one pulse per pitch period only. This pulse is variable in form and in amplitude, but marks the beginning of pitch periods and is further shaped in subsequent stages.

Block 8: This is the slicer stage which blocks the transmission path for pulses of insufficient amplitude,at the same time limiting their peak value. Effectively, therefore, a slice is taken through the pulses applied to this stage, producing an output which is either zero or of fixed amplitude. There is a small transitional region in which operation is uncertain, but in practice its effect may be ignored, since it occurs at a level which would not be present once the signal gate (Block 4) has opened. Although of constant amplitude, the width and shape of these pulses are still variable.

Block 9: This is a Schmitt gate which shapes the output of Block 8 into short pulses with fixed rise time and amplitude. It has the additional function of serving as a gate which is controlled by the voiced-unvoiced sound detector (blocks 15 to 22). It produces output pulses only on voiced sound in order that undesired pulses shall not be transmitted to subsequent stages. Its output is differentiated and the ensuing short pulses are used to trigger a monostable multivibrator via a diode which passes only pulses of negative polarity.

Block 10: The function of this monostable multivibrator is to form pulses of fixed duration, shape and amplitude and which possess sufficient energy to operate the electronic switches of the next stage.

Block 11: The strobing pulses of Block 10 are used here to start, stop and reset a timing device which produces an output which is a function of the pitch period. (It is approximately proportional to the logarithm of the reciprocal of the pitch period.) The principle of operation of this device is described on page 1068 of Ref. 1 and full details of its design are given on page 26 of this thesis. The pitch analogue is obtained from a cathode follower forming the output stage of this section.

Block 12: This is merely an amplifying stage serving at the same time to match the low impedance of the deflection coil of the cathode ray tube. A power transistor is used as an emitter follower here.

Block 13: The display unit. In view of the fact that the rotating cathode ray tube of the experimental model, although cumbersome to construct, produced a very acceptable type of display, this system was retained. Modifications were made as follows:

(a) A cathode ray tube of 12" diameter was used instead of the 9" tube, in order to obtain a larger viewing aperture.

(b) The screen type selected has image storage properties which are more favourable than the usual long persistence screen.

(c) The design of the mechanical drive system and the electrical slip-ring connections was modified to ensure smoother operation and less mechanical noise.

(d) Since the new type of screen allows a longer viewing time, it was arranged that the drive could be stopped so that the patterns "written" onto the screen can be studied at leisure.

Blocks 15 to 21: These eight blocks perform the function of voiced/unvoiced sound discrimination, their output preventing the display of a pattern when anything but voiced sound falls on the microphone. As explained before, the criterion used to differentiate between the two types of sound is the distribution of energy over the frequency spectrum. Since voice pitch determination takes place on the signal emerging from the 90 to 720 Hz filter, the simultaneous presence of strong components within this range, and weak components in the range above 4 KHz, almost invariably indicates the presence of a voiced sound. Hence this signal is applied to a rectifier producing unidirectional negative-going pulses (Blocks 17 and 18).

When frequency components above 4 KHz are large relative to those in the 90 to 720 Hz range, the sound is likely to be an unvoiced one. If large components below 50 Hz are present a plosive (like poor) is indicated. For both these classes of sound, the display should be suppressed. This is done by applying the full-band signal to the high-pass filter of Block 15 (fo = 4 KHz) and simultaneously to Block 20, which is a low-pass filter (fo = 50 Hz). Both these are simple active RC filters, having 24 dB per octave slope beyond cut-off. The output of these two filters are combined in Block 16, which allows attenuation of the one relative to the other in order to obtain a suitable balance between low and high frequency components. This summation signal is then amplified (Block 21) and rectified (Block 22) to obtain positive-going unidirectional pulses. The two inputs to Block 19, being of opposite polarity, are subtracted from each other when combined in this resistive network. Provision is again made to attenuate one relative to the other.

After smoothing, a control signal is obtained which changes polarity according to the nature of the sound: on silence it is zero; on voiced sound negative and on unvoiced sound positive. The controlled stages are arranged to pass information only if this control voltage is more negative than a preset threshold. The circuits which are subjected to this control, are the Schmitt gate (Block 9) and the C.R.T. intensity control electrode.

Finally there is an unvoiced-sound indicator (Block 14) which operates when this control voltage is sufficiently positive.

— — — oOo — — —

CHAPTER 5.

SHORTCOMINGS IN THE INSTRUMENT BUILT ACCORDING TO THE ORIGINAL DESIGN.

The reader should now be familiar with the principle of operation, as introduced in Chapter 4. More specific details were not presented at that stage, since the final design departs from that originally envisaged. It is felt that the reasons for these departures should be dealt with first, in order that details which were discarded in the final design, should not receive as much consideration as others which were finally adopted.

After the usual technical difficulties had been eliminated some basic problems emerged which will now be described.

1. ERRORS IN PITCH ANALOGUE: Since this is of primary importance and since the system developed to recognize errors and to eliminate these from the display have a bearing on other aspects of the design, this problem will be discussed first.

(a) Origin and form of Errors: It has been mentioned (see pp. 5 - 8) that the method by which the human sense of hearing assesses the pitch of a sound, is probably based on a continuous frequency spectrum analysis, the results of which are then further processed by the brain. Undoubtedly memory and previous experience play a role in the final assessment performed by the brain, which is able to discard irrellevant and erroneous information. To build an instrument which can imitate this function, is obviously a formidable task, as has been pointed out (page 7) and a simpler approach which yields acceptable results for the present purpose is sought. At the same time it has to be accepted that this simplified system will introduce errors which do not occur in the case of the sense of hearing and it will be necessary to devise means to detect and eliminate the more serious of these errors before displaying the pitch information.

Of the two simplified approaches available, the "waveform approach" has been selected for the present application, since the other method, based upon detecting the regularity of the waveshape in large sections of speech, requires the use of manually selected filters which cannot be tolerated. The selected approach operates on a cycle to cycle basis and is aimed at obtaining a pulse series in which each pulse represents one cycle in the movement of the vocal cords.

There are two potential sources of error inherent in this system:

(i) the inevitable presence of the vocal passage with its frequency-sensitive transfer characteristic between the vocal cords and the sensor (the Microphone) modifies the shape of the original glottal pulses in such a way that it becomes difficult to recognize them.

(ii) The vocal cords, being aperiodic, can change from cycle to cycle the rate at which the air stream from the lungs is interrupted. Pulses may be added to or omitted from a pulse train which, for its greater part, consists of regularly recurring pulses. These effects are smoothed out by the finite Q of the vocal cavities and subsequently ignored by the brain, but when analysis is carried out on a cycle to cycle basis, as in the present instrument, they lead to serious errors which the eye cannot be expected to ignore.

Errors occur frequently, since speech is a continuous succession of transients and the glottal pulse frequency often changes in a random fashion at the beginning and end of a sound or at the transition from one sound to the next. Fig. 2 illustrates a few of these errors. The waveshape of the microphone signal produced by a male speaker saying the words "sea shore" is shown in Fig. 2 a, whilst Fig. 2 b shows the pulse output of the shaper stages. Note the irregular pulse spacing at the beginning and end of both vowel sounds, and the great distraction which these cause in the pitch display, as shown in Fig. 2 c. Note also that the central portions of both vowel sounds produce acceptable pitch displays, the characteristic rising pitch on "sea" and falling pitch on "shore" being clearly portrayed.

(b) Detection and Elimination of Errors: It is evident from Fig. 2 c, that gross errors are those which cause large discontinuities in the pitch display. These can be in the form of a sudden rise or a sudden fall in the displayed pitch. If these are eliminated, as in Fig. 2 d, the eye is satisfied that the display is a reasonable replica of the pitch pattern as heard. Since the natural rate of change of pitch is limited (even though rates of 5 to 10 octaves per second occur, notably during the voicing of some dipthongs) excessive changes in periodicity from pulse to pulse may be used as an indication of the presence of spurious pulses.

Hence it is only necessary to determine the rate of change of pitch and if this exceeds a predetermined amount, to form a control signal which will suppress the display until such time as the pitch is again changing at a sufficiently slow rate. Admittedly this expedient will suppress portions of the display which occur at moments when the sense of hearing is able to assess the pitch by virtue of the more sophisticated method used by it. However, the resultant display is quite acceptable to the sense of sight which by its nature is not able to measure the passage of time with the accuracy of the sense of hearing. Consequently, relative to their aural duration, shortening of the visible patterns is not very evident. Furthermore, when a continuous line in the display becomes segmented, due to the action of the error suppressor, the eye is still able to follow the shape of the pitch curve quite well.

(c) Effect on other Parts of the Instrument: The pulse gate (Block 9 of Fig. 1) was now no longer necessary, since even if the period measuring stages are fed with erroneous pulses, these will in general be of random timing and will, therefore, bring the error suppressor into action, which will prevent their display. Fig. 3 shows that, when the pulse gate is removed, spurious pulses, although present in Fig. 3 b, do not appear in the pitch display of Fig. 3 d.

There is a temptation to consider discarding also the voiced-unvoiced circuits and allowing the error suppressor to take over this function. However, this is not possible, since the display has to be suppressed during periods of silence when the pitch rate-of-change is obviously zero and no output is produced by the error suppressor. Furthermore, the unvoiced-sound indicator has to be controlled by the voiced-unvoiced sound detector. For the record, the circuit of the discarded pulse gate is shown in the Appendix, Fig. A1.

2. MALFUNCTIONING OF VOICED-UNVOICED DETECTOR:

The proposed method of operation of this part was explained on pages 11 to 12. In practice it was found, however, that the output of the low-pass filter (intended to represent the soft plosives) had a very unfavourable signal-to-noise ratio making it almost impossible to detect the presence of these very-low-frequency components. The reasons for this are:

(a) the energy content of the soft plosives is relatively low;

(b) low frequency components of large amplitude are usually present in the ambient noise - often due to breath flutter when speaking close to the microphone;

(c) the response of some microphones to frequencies below 50 Hz is so low, that circuit noise becomes important.

Fortunately the failure of the low frequency portion of the unvoiced-sound detector to produce satisfactory results, is not of great importance, since the soft plosives, which were supposed to be detected, are of a transitional nature and give rise to a type of erroneous pitch output which is readily recognized by the error detector and consequently suppressed in the display. Block 16 and Block 20 of Fig. 1 could therefore be omitted in the final design. Admittedly the unvoiced-sound indicator does not respond to soft plosives, but it hardly seems feasible to devise a simple system of achieving this objective which, in any event, is of such small consequence that it was not pursued further. The circuitry of Block 16 and Block 20 is, however, shown in the appendix in the interest of completeness (Fig. A2).

3. POOR SIGNAL-TO-NOISE RATIO DUE TO THE ACTION OF THE AUTOMATIC GAIN CONTROL SYSTEM:

The intended function of the gain controlled stage preceding the filter stages (Blocks 2 and 3) was to provide an output whose dynamic range was small compared with that of the microphone signal. The purpose of this was to ensure that the shaper stages would not be subjected to a signal, the dynamic range of which exceeded the limits of their reliable operating range. The latter is determined by non-linearities which occur in the shaping diodes at low levels and in the buffer amplifiers between shaper stages at high levels. The maximum range acceptable to the shaper stages was estimated to be 40 dB.

This range was considered to be insufficient to accommodate all voice amplitudes. The addition of a gain controlled stage extended the range to over 60 dB, but introduced a new difficulty.

Under practical conditions the acoustic signal-to-noise ratio seldom exceeds 60 dB and hence when speech stops and A.G.C. increases the amplifier gain, room noise was amplified to almost the same level as the desired voice. This necessitated the incorporation of a gate which was controlled by the A.G.C. voltage in such a manner that it only passed the signal when its amplitude exceeded a certain preset level. When this is done however, it is found that the range between the level at which the gate first passes the signal and the maximum level which the A.G.C. controlled stage can accommodate without serious distortion, is not much more than 40 dB.

The situation may be summarized by noting that, by definition the ratio: Max.-signal/min.-signal is smaller at the output of a gain-controlled amplifier than at its input. Since the minimum signal in this case is noise only (be it ambient room noise or electrical noise in the amplifier) this ratio becomes: max.-signal/noise and it follows that this ratio must be smaller at the output of the amplifier than at its input. If a signal gate is used to transmit the signal only when it is above a certain threshold, the minimum output signal will be zero, but the dynamic range of input signals which will allow the gate to remain open, will then be much smaller than that which the amplifier can handle without the gate.

It can be said, therefore, that if the signal-to-noise ratio of the acoustic signal is not substantially larger than the dynamic range of the circuit to which it is applied, then no purpose is served by incorporating an A.G.C. stage and a signal gate.

Hence these stages were eliminated in the final design, but are shown in the appendix (Fig. A3).

--- oOo ---

CHAPTER 6.

ADDITIONS AND MODIFICATIONS TO THE ORIGINAL DESIGN.

Having attended to the basic shortcomings mentioned in Chapter 5, the instrument was subjected to practical trials by the staff and scholars of the School for the Deaf, Worcester, resulting in some further modifications and additions being affected. These will now be described.

1. PROVISION TO DISPLAY THE AMPLITUDE FUNCTION:

The concept of pitch is one which a completely deaf child finds difficult to grasp. Differences between a loud and a soft voice are more elementary and are associated with tactile sensations in the organs of speech in a more direct way than are differences in pitch. Since correct rhythm in speaking and correct control over the volume of the voice are perhaps just as important as proper control over voice pitch, it was deemed advisable to provide facilities in the instrument to display patterns of amplitude vs. time, in addition to the normal pitch vs. time patterns. Selection is by means of a front panel switch which connects the deflection amplifier of the C.R.T. either to the period analogue circuits or to the output of the linear amplifier preceding the filters. At the same time spot position and brightness are changed by this selector to put a continuously visible trace along the middle of the viewing aperture in the case of the amplitude display. When switched to produce pitch patterns, the spot is shifted to a position just below the bottom right-hand corner of the aperture and is extinguished until a voiced sound falls on the microphone. The amplitude function is therefore displayed in the form of a wave envelope pattern.(see Fig. 4) and the pitch function as a normal line graph. In both cases the time base is provided by the rotating cathode ray tube which portrays a persistent image moving across the viewing aperture from right to left.

2. ADDITION OF A TACTILE UNIT:

It was realized from the outset that a person with normal hearing finds the pitch or amplitude patterns meaningful after a very short period of practice, simply because of the fact that association between the senses of sight and of hearing is for him an everyday and automatic process. Since this process is impossible in the case of a deaf person, it was hoped that lip movements and C.R.O. patterns could be correlated by the child, and to this end a mirror was provided immediately below the viewing aperture in order that the child may see his own face and that of the teacher whilst at the same time watching the patterns.

Although this mirror is found very useful as a link betwee child and teacher, it soon became clear that a separate stimulus was required which did not involve the sense of sight, in order that the latter may be devoted entirely to the study of patterns on the screen. The sense of touch was the obvious choice, since deaf children are taught from an early age to feel the vibrations of the chest, throat and nose by means of their fingertips.

The necessary power amplifier to drive an electro-mechanical vibrator of the moving iron type was therefore provided as part of the complete instrument. In cases where a useful remnant of hearing is present, earphones may be used instead, if preferred.

3. <u>MANUAL SELECTION OF PITCH RANGE</u>: It has been shown (Ref. 1, pg. 1067) that the design of the shaper stages has to be a compromise. The effect of large secondary peaks in the waveshape produced by a male voice, has been illustrated and discussed on page 14. Indeed, it has been shown that the method of finding the pitch differs basically from the process performed by the human sense of hearing and that errors are therefore only to be expected. (See pages 5 to 8). By suitable design it is possible to minimize these errors when certain restrictions are stipulated for the input signals, such as limited pitch range, high signal-to-noise ratio, etc. It is to be expected, therefore, that when the design is optimized to deal with, for instance, low pitched male voices, in as much as it is rendered insensitive to secondary peaks, then it will no longer be optimum for high pitched voices, since it then tends to regard rapid amplitude fluctuations, such as those caused by "breathyness" on the voice (to quote one example) as the basic pitch, and proceeds to extract this information.

The design of the shaper stages is dealt with more fully in Chapter 8, but the above should suffice to show that it is necessary also to compromise on the design objective which aims at reducing the number of controls to the minimum . It was found that fewer errors occurred when the 90 to 720 Hz range was split into two overlapping bands which may be selected manually.

Practical experience shows that this switch presents no great hardship to the users, since many of the teachers' and children's voices fall within the same pitch range. The switch may therefore be left in one position when changing from child to teacher, and even if the incorrect range is selected for the voice concerned, the display is usually still meaningful, even though it then contains more blanks (corresponding to suppressed errors) than would have occurred had the correct range been selected.

In the final design, the pitch range selector was mechanically combined with the pitch/amplitude selector. This control therefore now has three positions: VOLUME, LOW PITCH and HIGH PITCH, respectively.

4. PROVISION OF MANUALLY TRIGGERED NON-REPETITIVE TIME BASE INSTEAD OF ROTATING C.R.T.:

There are obviously severe disadvantages in using a revolving cathode ray tube in order to provide a time axis. The following facts may be mentioned:

(a) a large C.R.T. has to be used in order to get a convenient size of display;

(b) mounting this large and fragile tube in a rotating structure and providing electrical connections to it via sliprings, is a cumbersome and expensive procedure;

(c) the time base is in the form of an arc of a circle instead of the customary straight line.

In view of the fact that reproduction of the instrument is considered in order to make it available for general use also at other schools for the deaf, it was deemed advisable to investigate the possibility of providing a simpler form of time axis which could be tried out by the scholars. After becoming acquainted with the use of both, the relative advantages may be weighed to establish whether the continuous time axis is justified in view of the extra cost and complexity involved in providing it.

In the case of the non-repetitive time base, the C.R.T. spot appears at the left-hand side of the screen when triggered by depressing a button on the microphone. The spot thereafter proceeds horizontally across the screen for as long as the button is held, or until it disappears at the right-hand side. The speed of the spot may be preset to any desired value - one inch per second appears to be satisfactory.

Fig. 5 shows pitch patterns of the same utterance displayed on the continuous, curved time base and on the linear time base respectively. The subjective effect of a moving spot, leaving behind it a persistent image, cannot be illustrated here, but is possibly more acceptable than the reversed procedure associated with the moving C.R.T. in which the medium on which the persistent image is formed, moves under the writing beam. It can be said that it seems more natural for past events to fade from view, remaining stationary as they do so, as in the case of the triggered time base, than to move away out of sight behind a mask, as in the rotating display.

It does not necessarily follow, of course, that someone whose mind is completely unconditioned, as far as graphical presentation of time sequencial events is concerned (such as a deaf child) would share this view or the conviction that a cartesian co-ordinate system with rectlinear axes is more acceptable than a circular display. These are factors which can only be determined experimentally and which are dealt with in Chapter 10.

- - - oOo - - -

CHAPTER 7.

RESUMÉ OF FINAL DESIGN.

Before proceeding with details of the design, a summary will be given of the final prototype as it exists in its present working form. This is done with reference to the block diagram of Fig. 6.

1. FUNCTION SELECTOR IN THE "PITCH" POSITIONS:

The microphone signal is applied to a preamplifier which feeds a phase splitter A. This in turn drives the power amplifier B to which earphones or a tactile unit (vibrator) may be connected. One of the two antiphase outputs of A is selected and applied to buffer amplifier C, which has a low output impedance to drive the following stages, filters D and G.

The amplifiers E and H receive their inputs from these two filters respectively. Rectification of their output signals by stages F and I produce positive-going and negative-going pulses respectively. These are suitably weighted and added in stage J to produce, after smoothing, a summation signal which has a positive polarity when high frequency components predominate in the microphone signal (as on unvoiced sound) and a negative polarity when a vowel sound is present.

This control signal is applied to the indicator K, causing it to glow on unvoiced sounds. At the same time the control voltage is also applied to the cathode of the C.R.T., thus turning the electron beam on or off, as required by the nature of the sound falling on the microphone.

The output of the band-pass filter G is also applied to the two cascaded shaper stages L, the output of which feeds the third shaper M when a low pitched voice is to be portrayed. When dealing with a high pitched voice, the function selector is placed in the appropriate position, which then eliminates the first two shaper stages.

The waveshape emerging from the shapers is normally in the form of a single pulse per pitch period. This enters the slicer N where low level noise components etc. are removed and the top of the pulse is clipped. Thereafter a Schmitt switch O further forms these pulses to uniform dimensions and thereafter applies these as triggering pulses to the sampling pulse generator P.

The function of P is to generate pulses of constant width which are sufficiently powerful to operate the timing device Q and its associated information storage system. The output of Q is taken off via a cathode follower, since the latter imposes the required negligible load on Q. The cathode follower drives the C.R.T. deflection amplifier R, a transistor having a fairly low input impedance. The resting position of the spot is behind the lower right-hand edge of the aperture in the mask over the C.R.T.

The pitch rate-of-change detector S derives its input from the cathode follower in the deflection amplifier. Its output is in the form of a short pulse each time the rate-of-change of

---

* The purpose of the polarity selector is to provide means to ensure that the initial voice peaks which mark the beginning of pitch periods will be of the correct polarity for proper operation of the shapers, whatever the polarity of the microphone.

pitch period from one cycle to the next exceeds a predetermined amount. Each of these pulses triggers a monostable multivibrator to produce one output pulse of constant form and duration which is then further shaped by two diodes and an R.C. network T, to produce blanking pulses. These blanking pulses serve to extinguish the trace for the duration of false information which originates in the pitch extracting circuits.

2. FUNCTION SELECTOR IN "VOLUME" POSITION:

In the case of the amplitude function display, the cathode follower in R obtains its input from the linear amplifier at A. At the same time the D.C. component of the input to R is altered so as to deflect the stationary spot to the middle of the viewing aperture. Other sections of the selector switch (Sc and Sd) are responsible for appropriate changes in spot intensity as required by the two types of display.

Not shown on Fig. 6, are details of the mechanical construction of the instrument which, however, does not differ in principle from the instrument described in Reference (1). The special cathode ray tube used calls for infra-red illumination to render the stored image visible and provision for an infra-red source had to be made.

The power supply incorporates a screened ring-core transformer to minimize external magnetic fields which would give rise to spurious deflection of the C.R.T. spot, since the transformer is situated close to the neck of the tube. The E.H.T. requirements of the tube are met by an R.F. type generator supplying the required 12 kilovolts. All component parts are housed in one enclosure 15" x 15" x 22" long, as shown in Figs. 29 - 33 and 35.

--- oOo ---

CHAPTER 8.

DETAILS OF FINAL DESIGN : ELECTRICAL PART.

There are obviously some portions of the instrument which consist of perfectly straightforward electronic circuitry and little purpose will be served in documenting the detailed design of these. In cases like these, reasons will merely be quoted for the particular methods used.

It is, however, the writer's belief that certain other portions of the instrument represent a novel approach and these will be covered in more detail. The instrument will be described in its role as pitch indicator, with the function selecting switch set to "Low Pitch". The letters used, refer to the block diagram of Fig. 6. It is recommended that the circuit diagram (Fig. 7) be consulted in conjunction with this description.

Block A : Preamplifier and Phase Splitter V1:

Since the microphone is used close to the speaker's lips in the interest of obtaining the best possible acoustic signal-to-noise ratio, the input signal is large (typically 2 to 200 millivolts R.M.S.) and consequently no extensive precautions need be taken to minimise first-stage noise. One half of an ECC83 was used here with contact potential bias. It feeds the other half of the same valve (used as a long-tailed phase splitter) via a preset volume control. The outputs of the phase splitter are taken to the microphone polarity selector and also to Block B.

Block B : Power Amplifier V2:

The purpose of this stage is to drive earphones or the tactile unit, the latter requiring several watts. Using an ECL82 valve, it was possible to apply negative feedback between the secondary of the output transformer and the input circuit of the triode section of this valve. Thus good linearity is ensured at levels up to 2 watts as shown in Fig. 8. When using the tactile unit, considerable distortion of the wave-shape can be tolerated and 4 to 5 watts obtained.

Block C : Buffer Amplifier V3a:

The main purpose of this is to provide a low impedance feedpoint for the two filters which follow. It also provides some gain. The valve used here is ½ ECC81, the other half being used in Block J.

Block D : High-pass Filter V16:

The function of this is simply to pass signals in a frequency range where the energy content of voiced sound is low and that of unvoiced sound relatively high. A steep cut-off is not required and various writers have suggested cut-off frequencies of between 3500 and 4500 Hz. (See pg. 59 of Ref.6 and pg491 of Ref. 16.) Typical spectrograms of some sustained vowels and consonants are shown in Fig. 9 (male voice) and in Fig. 10 (female voice). These spectrograms were obtained by recording the sound on an endless loop of magnetic tape of about two seconds duration, this being repeatedly played back via an octave filter into a voltmeter.

The ordinates of these graphs represent the output obtained from the filter at the different octave bands, the centre frequencies of which covered the range 112 to 9600 Hz in half-octave steps. It was endeavoured to keep the voice pitch the same for all vowels and the pitch used, was chosen to fall at the centre of one of the octave ranges. The entire equipment, including the microphone, tape recorder, filter and voltmeter, was calibrated before the test started and the system gain was adjusted to place the peak of each sound spectrogram at the same reference level (odB).

The vertical lines at 700 Hz on these graphs represent the upper cut-off frequency of the band-pass filter and it is immediately evident that a large fraction of the total energy of voiced sound lies in the region below this frequency. This applies equally to male and female voices.

The problem is to decide on a suitable frequency at which to place a vertical line representing the cut-off frequency of the high-pass filter. Since there are relatively strong components of voiced sound up to at least 6 KHz, and the amplitude of unvoiced sound components appear still to be rising beyond 9 KHz, there is the temptation to stipulate a cut-off frequency for the high-pass filter even higher than 6 KHz. It should be remembered, however, that the above spectrograms represent an overall flat frequency response up to at least 10 KHz, whilst in practice the microphone used,may not have a flat response at frequencies as high as these and also poor signal-to-noise conditions may prevail beyond 6 KHz.

A reasonable compromise seems to be to place the cut-off frequency at 4 KHz. In the worst cases voiced components above 4 KHz are still 20 dB below their peaks in the 90 to 720 Hz range and the strongest components of unvoiced sound in the 90 to 720 Hz range are 12 dB below peak components in the range above 4 KHz for both male and female voices. There is thus a fair margin of safety, provided the relative levels of the two frequency bands are correctly set, as will be described under Block J.

The filters of Block D and Block G are of the active type, using frequency dependent R-C networks in the feedback loops around cathode follower gain stages (17). By cascading two such stages, a slope beyond cut-off of 24 dB per octave is obtained, which is entirely adequate for the present application. Of the many different types of active R-C filters which have been described, this one was chosen because of its inherent linearity and large signal handling capabilities (being a cathode follower). Details of the design are given in the Appendix page Ai . The two sections of an ECC81 valve are used in this stage and its response is shown in Fig. 11.

Block E : Amplifier V15a:

The filter feeds this Class A amplifier which raises the signal to a level sufficient to drive the next stage. One half of an ECC82 valve is used here and a 100K ohm stopper is included in the grid circuit in order that grid current, flowing during the positive peaks of large amplitude signals, should not load the filter and affect its characteristics.

Block F : Rectifier D20, D21:

This is a voltage doubler circuit, using two silicon diodes having a 400 v inverse rating each. With the value of supply voltage and load resistance used, the peak-to-peak output of the preceding stage can never exceed 250 volts. These rectifiers are therefore adequate. A voltage doubler was used since (i) it is isolated from the D.C. level of the stage feeding it by means of a coupling condenser thus eliminating the need for using transformer coupling, and (ii) it provides a larger rectified output than that of a half-wave type. The rectifiers are connected to produce an output of positive polarity.

Block G : Band-pass Filter V4, V5:

The detailed design of this is given in the appendix as for the high-pass filter of Block D. (See page Ai). Here pairs of high-pass stages (fo = 90 Hz) are cascaded with pairs of low-pass stages (fo = 720 Hz) to form a 90 to 720 Hz band-pass filter, once again having an attenuation beyond cut-off of 24 dB per octave. In all, two ECC81 valves are used. Fig. 12 shows the response of this stage.

Block H : Amplifier V15b:

This is identical to Block E and uses the other half of the ECC82 valve.

Block I : Rectifier D18, D19:

This is identical to Block F, except that the polarity of the rectifiers is reversed in order to produce a negative-going output.

Block J : Smoothing and Combining Stage. V3b:

The outputs of the two rectifying stages are smoothed and combined in suitable proportion in this stage to produce a negative-going signal for voiced sounds and a positive-going signal for unvoiced sounds. Since the polarities are opposite, this effect is achieved by simply connecting the output from the two rectifiers to the outer ends of a resistive potentiometer, the slider of which then assumes a voltage which is the sum of fractions of the two input voltages. The value of these fractions depends on the setting of the slider and was determined empirically.

The attack and decay time constants of the smoothing circuit used here have to be kept fairly short. A rapid attack is necessary in order to ensure that even short voiced sounds will be displayed. A short decay time prevents the display of spurious information after a voiced sound has stopped. On the other hand, too short a decay time will lead to insufficient smoothing in the case of a low pitched male voice and consequently to an intensity ripple in the display. On account of the low time-base speed, the latter effect is, however, only visible as a reduction in the intensity. A decay time-constant of 3 milliseconds was found to be a good practical compromise and was used in the case of both rectifying systems.

The potentiometer arm, which is a high impedance source, feeds the grid of a cathode follower V3b ($\frac{1}{2}$ ECC81). When V3b anode current is cut off by a large signal on its grid, the cathode voltage drops to about -20 volts since the cathode is connected to a point between the -16V and the -105V supplies via 15K, 33K and 150K resistors. (C.R.T. beam current has also to be taken into account when estimating this voltage). Under no circumstances can this voltage be exceeded. Since the intensity of the cathode ray spot is increased by this negative control voltage (applied to the C.R.T. cathode) and since voiced sound produces cut-off bias on V3b on all but the very weakest of sounds, the C.R.T. spot is in practice either extinguished or at full brightness. The transitional region of below-normal brightness is small and occurs over a very limited range of voice intensity. At the same time the C.R.T. intensity is limited since the control voltage can never exceed this fixed value of -20 volts. Hence screen burning is avoided.

Positive excursions of the control voltage are not so important, as far as the C.R.T. is concerned. However, the unvoiced-sound indicator should be protected from these as described under Block K.

It should be pointed out that such limiting of the control voltage can only be done after the rectified outputs of the band-pass and high-pass filters have been combined, since if appreciable amplitude limiting were allowed to make place in the filter stages or subsequent amplifier stages, the energy distribution depicted by Figures 9 and 10 would be flattened and this would make it almost impossible to distinguish voiced from unvoiced sounds.

<u>Block K : Unvoiced Sound Indicator. V18 (DM70)</u>:

This should be extinguished on silence and on voiced sound and should glow on unvoiced sound. H.T. and filament circuits were proportioned according to the recommendations of the data sheet of this valve type and the control grid was supplied from a potentiometer connected between the cathode follower V3b in Block J and the -105 volt line in order to establish the correct D.C. potentials for this stage. Since excessive voltage excursions in the positive direction of the control grid cause "blooming" of the green glow of the DM70, such excursions were limited to less than about +3 volts by die diode D17, which is returned to a +3 volt point in the circuit. A certain amount of smoothing is applied to the signal reaching the grid by the 0.1 MF condenser to prevent rapid flashing on transients. A time constant of about 3 milliseconds was found to be adequate here.

Typical grid voltages are as follows:

Voiced sound	:	-34
Silence	:	-17
Unvoiced sound	:	+ 3

Blocks L and M: Shaper Stages 1, 2 and 3. V6:

The principle of operation is the same for all three shaper stages and has been fully described in Reference 1, page 1067. However, in the final design, a further property of the speech wave is exploited in addition to the two properties on which shaping was based initially (first peak in pitch period usually the largest and slope associated with this peak usually the greatest). This third property is that the area under the first peak is usually larger than areas under subsequent peaks in the pitch period.

Integration of the waveform produced by the circuits which recognize: (a) maximum slope and (b) maximum amplitude, will produce an output which is a function of the area under the waveform. Integration has the further desirable effect of smoothing out ripples caused by large amplitude mid-frequency harmonics often present in low-pitched male voices.

Fig. 13(a) is an oscillogram of the output of the band-pass filter for the sound u as in pool as produced by a male speaker with a low-pitched voice. Inspection of the waveshape reveals that only weak higher harmonics are present, a fact which is confirmed by the data of Fig. 9. However, prominent lower harmonics cause a secondary peak to appear on the waveform as shown in Fig. 13 (a).

Fig. 13 (b) shows the result of passing this signal through a shaper stage consisting of a slope detector (differentiating circuit) followed by a peak detector. Note that the secondary peak is not suppressed by the shaper, although its amplitude relative to the main peak is reduced.

Fig. 13 (c) illustrates the effect of applying a stage of integration to this waveform. Here the relative amplitude of the secondary peak is further reduced.

The shaper circuit used in the original instrument was re-arranged somewhat so that the diode peak detector was placed between the differentiating and integrating components. At the same time this change resulted in the reduction of the total number of components, in spite of the addition of a stage of integration. Solid-state diodes were used instead of thermionic diodes as in the original instrument.

The new circuit is shown in Fig. 14. When the input signal swings positive, D conducts and the charging current of C_1 causes a voltage drop to appear across R3, which is a function of the rate-of-change of input voltage. After the peak of the wave, D no longer conducts and therefore the voltage across R3 falls to zero, where it remains until a sufficiently positive peak is applied to cause D to conduct again. The function of R_1 is similar to that of R_1 in the original circuit (Fig. 2a of Reference 1) serving to discharge C_1 slowly in order that a new pulse will appear across R3 at the peak of the next cycle. C_1 now also serves as a D.C. blocking condenser and no further components are required (as were in the original design) in order to accommodate the different D.C. levels.

The integrating portion consists of R_2 and C_1. The component values listed in Fig. 14 were determined empirically, using typical speech wave input signals and assuming that the generator resistance would be approximately 10K ohms.

The following table shows the ratio between the amplitude of a sine wave input signal and the amplitude of the output pulse.

Frequency, Hz	85	200	750
Ratio	19	15	25

(R_1 = 560k ohms)

As a result of this attenuation, each shaper stage has to be followed by an amplifier, half of an ECC81 being used for this purpose with careful attention to H.T. smoothing and optimum choice of operating parameters to ensure that the useful dynamic range will be as large as possible.

Since a single stage of amplification gives a phase reversal and since the assymmetrical speech wave has to be correctly polarized for proper operation of the shaper stage, the diode of the second shaper is reversed relative to those of the first and third shapers. In all other respects the three shapers are identical.

Block N: Slicer. V7:

The signal delivered by the shapers is normally in the form of one pulse per pitch period. This pulse is, however, variable in shape, in amplitude and, of course, in repetition frequency. In addition there are low-level noise components which leak through the shaper stages, forming a continuous background. These pulses are, therefore, sliced at a preset level which is above most of the background noise. For this purpose V7a (½ ECC83) is biassed beyond cut-off so that only input signals which exceed the threshold are reproduced, in amplified form, at the anode of this valve. The negative-going output pulses of this stage have an amplitude far in excess of the grid base of the second ½ ECC83, V7b, to which they are applied. Consequently only that portion of the pulse which lies near the zero line appears in the anode circuit of V7b. The total effect of V7 is, therefore, to take a slice through the signal applied to it.

The amplitude of the output pulses is reasonably constant, but their rise time, although much shorter than that of the original pulses, is still variable and so is the duration of the pulses. Since the aim is to provide short pulses of constant amplitude to the strobing circuits of Block P, it was necessary to incorporate a further stage of pulse shaping.

Block O: Schmitt Switch. V8:

This stage has two stable states. It switches from one state to the other at a critical input voltage and then remains in that state irrespective of further excursions of input voltage beyond that point. It switches back to the original state at another critical voltage. The leading and trailing edges of its output pulses, therefore, occur at well-defined values of the input wave and their rise- and fall-times are short and are determined not by the input waveshape, but by time constant of the stage.

The design of a Schmitt switch is adequately covered in the literature and will not be detailed here. See for instance (18)

The output of this stage is a pulse of constant rise time and amplitude. From this a short triggering pulse is developed by differentiation, and subsequent selection of the negative-going pulse is provided by a diode D4.

Note: The use of a triode-pentode valve (6U8 or ECF82) in this stage is not necessary, but in the original version of this design this valve also served as a pulse gate which was controlled by the voiced-unvoiced sound detector (Block J etc.) When this gate was found unnecessary (see page 14) it was convenient to retain the same valve in this position.

Blocks P and Q: Sampling Pulse Generator and Period Measuring Device. V9 to V13, D5 to D9:

The principle of operation of the pitch period measuring device was described in Reference 1, pg. 1068. Fig. 6 of this paper is reproduced here as Fig. 15, since it will frequently be necessary to refer to this in the following description of practical details.

Consider the case of a sound of constant 360 Hz pitch which commences after a period of silence and is then maintained. Fig. 16a depicts the output of the sampling pulse generator, these pulses serving to close S1 (Fig. 15); Fig. 16b shows the time-delayed charging pulses which close S2, Fig. 16c the wave-shape present across the timing network R3.C3 and Fig. 16d the form of the output signal present across C4.

The first charging pulse closes S2 and, commencing from a completely discharged state, C3 charges to a voltage V. When S2 reopens, C3 discharges exponentially through R3 until S1 is closed by a sampling pulse. This places C3 in parallel with C4 and causes a small drop in the voltage present across C3, as shown, since the charge left in C3 is now shared by C3 and C4. These two condensers now discharge into R3 for the duration of the sampling pulse, at the end of which S1 reopens, leaving C4 isolated. C4 therefore retains a voltage which remains constant (provided leakage from C4 is negligible) until it gets another small increment in voltage during the next sampling pulse.

It should be noted that the output, Fig. 16d, is a function of the time elapsed between the end of a charging pulse and the beginning of a sampling pulse, this time being shorter than the pitch period by the time elapsed between the beginning of the sampling pulse and the end of the charging pulse. Ideally the pulses should be of negligible duration compared with the shortest pitch period in order that the output may be an accurate measure of the pitch period. Inspection of Fig. 17 reveals that at a pitch frequency of 720 Hz a relatively small portion of the period is left for timing purposes when the pitch period and the switching pulses have the proportions shown. Small variations in stability of pulse duration will have a marked effect on the output under these conditions.

A further ill effect of a long pulse duration is that C4 is carried along the C3R3 discharge curve for the duration of the sampling pulse and this, in effect, causes a ripple on the pitch analogue (i.e. the output voltage across C4). This ripple is of course worst for high pitch frequencies, since for these the slope of the C3R3 discharge curve, at the instant the sampling pulse occurs, has a maximum value.

However, as the duration of these pulses is decreased, so the current which has to be carried by S1 and S2 in order to charge C3 and C4 in the available time, increases. Practical considerations, therefore, make it necessary to accept a compromise which allows accuracy to be adequate and the ripple on the display inconsequential, for the purpose for which the instrument will be used but which, at the same time, makes possible the use of switching elements of reasonable size.

Good stability of pulse amplitude and form ensures that reasonable accuracy of the pitch analogue will be maintained. The design of the sampling pulse generator should be such that this stability is achieved.

The logical procedure to follow in the further design of the period measuring device seems to be to decide on reasonable values of C3 and C4, then to find a relationship between pulse duration and charging current and thereafter to check whether this charging current can be carried by electronic switches of reasonable dimensions. Finally these may then be designed and a value chosen for R3 to make the output as nearby as possible a logarithmic function of pitch period within the limits 90 to 720 Hz.

Choosing C4:

This condenser has to maintain a voltage which remains as constant as possible between successive pitch period pulses. Since the display is on a C.R.T., the accuracy of read-out is limited by the radial resolution of the tube. The ratio of total radial deflection (which will represent the 90 to 720 Hz pitch range) to spot diameter was measured to be about 50. Hence a constancy of the above voltage to 2% is entirely adequate, since this will cause the display to drop by only one spot diameter during the time of the longest pitch period. The latter is 1/90 second and it may be assumed that the total leakage present across C4 may be represented by a resistor of not less than 15 megohms. Substituting these values in the expression for the discharge of a condenser (see Appendix pg. Aii) gives a value to C4 of 0.04 MF.

Choosing C3:

C4 is placed in parallel with C3 via S1 by the sampling pulse and since C3 and C4 are not charged to the same voltage at the times of occurrence of the first few sampling pulses of a pulse train, there is a stepped build-up of voltage across C4, as shown in Fig. 16d. It is desirable for the pitch analogue to reach its final value within as few pitch periods as possible. Furthermore, if the steps exceed about 5% of the maximum value of the pitch analogue, the error detector will interpret these as false information and suppress the display (see pg. 32). The first step is, therefore, never displayed since it commences at 0 and proceeds to a value close to the ultimate. The first possible opportunity at which the pitch analogue can be displayed, is, therefore, at the moment the second step has taken place and this step has to be less than 5% of the voltage which will finally exist across C4. In the Appendix (pg. Aii) it is shown that under these conditions, the ratio C3/C4 has to exceed 18. Hence C3 was made 1 MF, which results in a third step of 1/7%, which is negligibly small.

Duration of Sampling Pulse:

The condenser C4 has to be charged from C3 during the sampling pulse via the electronic switch S1, the design of which is given on pg. 28 where it is shown that S1 has a resistance when conducting of about 200 ohms. If C4 is to charge to within 1% of the ultimate voltage during a pulse (a reasonable stipulation) then the pulse duration must be no less than 0.04 Msec. If the maximum potential to which C4 has to charge, is 6 volts, then the initial current will be 6/200 amp or 30 Ma.

Duration of Charging Pulse:

During this pulse the condenser C3 has to charge via switch S2, which is of a different design compared with S1 and which passes a fairly constant charging current. It may be assumed that a current of 50 Ma will be available and that C3 has to be charged to 10 volts. Since it = CV, t = CV/i. Therefore: $t = 10^{-6} \times 10/0.05 = 0.2$ millisecond.

Design of S2:

The following requirements have to be met:

(a) The one side of S2 may be earthed if S2 and V are interchanged (Fig. 15).

(b) When closed, the switch has to pass a current of at least 50 Ma into C3, until it is charged to a potential of, say, 10 volts. The potential reached by C3 must be the same for all charging pulses.

(c) When open, the switch should have an insulation resistance of no less than, say, ½ megohm.

(d) It has to conduct current in only one direction.

A suitable circuit to achieve these objectives is given in Fig. 18. The triode-connected pentode section of an ECL82 valve is biassed to -35 volts which ensures anode and g2 current cut-off. It is only allowed to conduct for the duration of the charging pulse which is applied to its control grid The resulting anode current flowing through C3, therefore charges it. The Zener diode Dz does not conduct before the potential across it is about 9 volts, but when it does, this potential hardly rises any further, since a small increase in potential above that at which reverse conduction starts in a Zener diode is accompanied by a large increase in current (typically 0.6 volt rise requires a 50 Ma rise in current). Hence the voltage to which C3 is charged, is always the same (approx. 9 volts). It should be noted that, in practice, 0.3 to 0.5 Msec. pulses (depending on pitch) are applied to the control grid instead of the calculated 0.2 Msec. to ensure that the voltage reached by the condenser is determined by the Zener diode and not by the precise value of the current passed by the valve. This ensures that deterioration in valve emission or small changes in the shape of the charging pulses will not affect the constancy of the potential reached by C3. It also makes possible the use of pulses which are less perfectly shaped than those shown in Fig. 18.

The pentode section of an ECL82 valve was connected as a triode (g2 strapped to anode) and anode plus g2 current measured when Va = 100 volts and Vgl = 0 volts. This was 60 Ma. and hence adequate for the purpose (see (b) above). The other conditions are also easily met.

Design of Switch S1:

The following requirements have to be met:

(a) Both sides of S1 should be isolated from the common ground. Leakage resistance to ground should exceed 15 Megohms.

(b) When open, the leakage resistance between the output terminal and any other terminal of S1 should exceed 15 Megohms, even when the maximum voltage difference encountered in use is present across S1 (about 10 volts).

(c) When closed, the resistance between input and output ends of S1 should be small in order that C4 will charge to within a few percent of the voltage present across C3 in the short time available during the sampling pulse. The resistance should be independent of the direction of current flow

(d) The output of S1 should contain no switching transients.

The principle of operation of a suitable device is explained with reference to Fig. 19. The diodes D1, D4 and D2, D3 form two parallel pairs which are reverse-biassed by the battery B, the potential of which exceeds the largest potential difference anticipated between the signal points 1 and 0 with the switch in the non-conducting state.

The transformer phasing is such that switching pulses appearing across the secondary winding overcome this reverse bias and cause all four diodes to conduct heavily for the duration of the pulse. In this state the diodes can conduct an external current from point 1 to point 0, and also in the opposite direction. If the forward voltage drop is the same for each of the four diodes, there will be no potential difference between points 1 and 0 and the transformer current will divide equally between the pair D1, D4 and the pair D2, D3.

The signal current flowing between points 1 and 0 should not exceed the current circulating in the circuit between transformer and diodes, since if it does, one pair of diodes (e.g. D1, D3) will cease conducting and the effective resistance between points 0 and 1 will rise abruptly due to a complete change in the equivalent circuit.

To prevent the introduction of switching transients, the following conditions should be satisfied:

(a) The bridge consisting of the four diodes, should be in balance during the switching pulse, i.e. the ratio of the P.D. across D1 to the P.D. across D4, should equal the ratio of the P.D. across D2 to the P.D. across D3.

(b) The two ends of the transformer secondary should have identical capacities to ground.

(c) There should be no capacitive coupling between primary and secondary windings, which would cause an in-phase voltage to appear at the two ends of the secondary.

The combination of four series-parallel connected diodes has the same forward resistance as a single diode, if it can be assumed that each of the four diodes carries the same current as this equivalent single diode and that all diodes are identical. Using two 6AL5 valves, this resistance is on the order of 200 ohms.

In the non-conducting state the points 1 and 0 remain isolated, provided the P.D. between them does not exceed the value of the reverse bias. The insulation resistance is that of one non-conducting diode and therefore quite adequate. Leakage between transformer winding and ground is easily made negligible; capacitive balance of the two halves of the secondary is readily achieved by placing the windings on two identical bobbins mounted side by side on the core. An electrostatic screen between primary and secondary prevents capacitive coupling. The design of the pulse transformer is dealt with in the Appendix page Aii1

In Fig. 19 a battery is shown providing the required reverse bias. The same purpose is served by the parallel connection of a Zener diode and a large condenser, as shown in Fig. 20. This operates as follows:

During a switching pulse, the current circulates between the transformer secondary and the diodes D1, D2, D3, D4 and charges C, the voltage rising until the reverse breakdown point of the Zener is reached and practically no further increase in voltage can take place. The entire circulating current is now diverted to the Zener. Therefore an adequate current is available throughout the pulse duration to ensure proper operation of the four-diode switch. At the end of the pulse, C only discharges rapidly until its P.D. has fallen to a value where the Zener diode reverts to its high resistance condition. Using two type OAZ 207 Zeners in series provides a reverse bias of about 17 volts with a leakage resistance of about 30 Megohms. Using a capacity of 0.25 MF for C, therefore results in a time constant of several seconds. Admittedly the charge across C will ultimately disappear and the switch will then no longer be open in the absence of switching pulses, but during sustained periods of silence the C.R.T. display is suppressed by the voiced-unvoiced circuit and therefore false responses, which may be produced by the switch in this condition, are not displayed.

The maximum P.D. which has to be withstood by the four-diode switch in its non-conducting state occurs at the instant of the first charging pulse of a train of pulses commencing after a period of silence. This P.D. is governed by the Zener diode which forms part of the charging circuit of the period-timing device (see page 28) and can never exceed 9 volts. A reverse bias, having an initial value of 17 volts and taking several seconds to decay to 9 volts, is therefore entirely adequate.

The four-diode switch has to pass a maximum current of 30 Ma. (see page 27) between its ends 1 and 0. This calls for a switching pulse current of at least 30 Ma from the transformer secondary, half of which passes through each diode. Since the instantaneous current rating for the 6AL5 is 54 Ma. per section this valve may be used with confidence.

Sampling Pulse Generator (Block P): V_{10}

With C (Fig. 20) charged to 17 volts and a current of, say, 50 Ma. circulating, there will be about 10 volts across the four-diode switch and a further drop across the equivalent resistance of the pulse transformer. A realistic value for the latter is 25 ohms (see Appendix page Aiv) which results in a further 1.2 volt drop. The total voltage around the circuit is therefore 28 volts at 50 Ma., which amounts to a peak pulse power of 1.4 watts. Allowing for losses in the transformer and for the fact that the pulse delay circuit is also fed from the primary of the pulse transformer, it is reasonable to provide for at least 2 watts of peak pulse power. This dictates the use of a valve in the power pentode class. In practice an ECL80 was used, the triode section completing the requirements of a monostable multivibrator.

Circuit constants are chosen to produce at the anode of the pentode section a negative-going pulse, the duration of which is about 0.1 Msec., thus ensuring that adequate time is available to charge C4 completely. The screen grid of the pentode section is connected to the 150 volt regulated line in the interest of obtaining pulses of stable height and duration. This low screen grid voltage limits the anode current to a value which, after step-up in the pulse transformer, will be within the maximum current limits of the diodes in the switch S1. A clamping diode across the primary of the pulse transformer eliminates the voltage surge which would, in its absence, occur at the end of the sampling pulse when anode current in the pentode is cut off abruptly.

Forming the delayed charging Pulses:

After the sampling process is over, the condenser C3 is again fully charged when the switch S2 closes. In practice S2 is closed by the application of a positive pulse to the grid of the charging valve (see Fig. 18) but this should take place only after S1 has opened again. The leading edge of the charging pulse should therefore occur slightly after the trailing edge of the sampling pulse. This function is performed by the circuit shown in Fig. 21. A sampling pulse causes D2 to conduct and the 0.005 MF condenser C1 to charge to the peak of the pulse, at the trailing edge of which the anode voltage of V10a rises, causing D2 to open and D1 to conduct, and the charge acquired by C1 is shared between it and C2. Hence a positive-going step appears across C2 at the instant that the trailing edge of the sampling pulse occurs. This charge is dissipated in the 33 Kohm resistor, which is chosen to produce the required duration of charging pulse applied to the control grid of V13b. The 270 Kohm limiting resistor in conjunction with valve input capacity ensures that there is a further small time delay and also that the grid voltage stays close to zero with respect to cathode for a controlled period of time.

Choosing R3:

A value has to be found for the timing network R3C3 (Fig.15) which will make the voltage present across C3 at the end of a pitch period as nearly as possible proportional to the logarithm of the pitch frequency. This voltage is a function of the time elapsed between the end of a charging pulse and the beginning of the following sampling pulse, and is less than the actual pitch period by the sum of the durations of these two pulses and the time delay between them. In practice this sum is approximately 0.4 Msec. A family of curves may now be drawn, showing the voltage retained by C3 vs. pitch frequency if the period corresponding to each frequency is reduced by 0.4 Msec. before being substituted in the usual equation for the discharge of a condenser into a resistor. The parameter of these curves is the product R3C3 and it is evident from Fig. 22 that when R3C3 5 milliseconds, the curve of E/V vs. log.f is the best possible straight line over the desired frequency range of 90 to 720 Hz.

Hence, since the value of C3 has already been fixed at 1 MF, R3 should be 5000 ohms, a value which is used in practice. Non-linearities in the deflection amplifier and in the C.R.T. itself have an adverse effect on linearity, but the final result is acceptable for the purpose as shown in Fig.23. See also Fig. 28 a and b.

Block R : Deflection Amplifier:

With the recommended 12 KV accellerating potential applied to the C.R.T. and using the deflection coil assembly supplied with the tube, the following measurements were performed:

Coil:	D.C. Resistance: ohms.	Current for 1" deflection:	Inductance mHy:
A	160	14 Ma.	Approx. 250
B	20	45 Ma.	Approx. 25

Fed from a constant voltage source, the inductance of the coil will cause the deflection to be frequency sensitive. In the case of Coil A, the response will be 3dB down at a frequency of approximately 100 Hz. This is quite sufficient to portray even the fastest change of pitch faithfully.

The maximum deflection required is about 5½", which calls for a current in Coil A of 77 Ma. at 10.8 volts, or, using Coil B, the required current is 248 Ma. at 5 volts. Since the use of battery power supplies are not considered and mains derived D.C. power supplies for low voltage, high current application call for the use of large ripple filtering components, it was decided to use Coil A. Since even this is a relatively low impedance device and constant voltage drive could be employed, which did not call for high drive voltages, the obvious choice of a deflection amplifier was a small power transistor. The supply voltage for this was conveniently obtained, using silicon diodes in a voltage doubling circuit, from the 6.3 volt filament supply. After smoothing, this produced about 16 volts.

Maximum dissipation in the transistor, operating under the above conditions can be shown to be 0.26 watt. An OC30 which has a safe dissipation of 1 watt can, therefore, be used with confidence. Used in the common collector configuration, the required drive current is only 1 to 2 Ma., which is readily supplied by a cathode follower V13a. The latter has to be incorporated in order to impose a negligible load on the four-diode switch which is connected to its grid. In order to establish the correct D.C. level at the base of the transistor so as to position the C.R.T. spot at the lower edge of the viewing mask, the cathode load resistor of V13a is returned to the -105 volt regulated line and the spot position control is made part of this cathode load resistor. Adjustment of this control obviously also results in a change of gain, but this is a second order effect only and once set, the position control needs no further adjustment.

The transistor is used in this configuration for the following reasons:

(a) Residual power supply ripple is greatly attenuated.

(b) Thermal stability is ensured.

(c) A good match to the thermionic valve driving the transistor is achieved.

(d) Correct operating potentials are more easily provided than in the common emitter configuration.

<u>Block S : Pitch Rate-of-Change Detector. D10 to D14:</u>

The pitch analogue appearing across the storage condenser following the four-diode switch (C4 in Fig. 15) changes in discrete steps. The size of these steps depends on the change in the period between successive pulses produced by the waveform shaping circuits. Hence, if the pitch of the input signal increases rapidly, these steps are large and negative-going. If there is a rapid decrease of pitch, the steps are large and positive-going. If these steps are differentiated, short pulses are produced, the amplitude of which is a function of the rate-of-change of pitch.

The requirements of the next stage are such (see pg.34) that the polarity of these pulses should be the same whatever the sign of the change in pitch. Hence a combination of a phase inverter producing pulses of both polarities and 4 diodes was used to ensure that all pulses appearing across the output load resistor, are negative-going.

The action of this circuit is explained with reference to Fig. 24, which shows the pitch analogue being applied to the input of a phase inverter. A positive step in the input causes a negative step in the anode circuit, which is conducted to R1 via D1. Since R1C1 is short (typically ½ Msec.) the voltage across R1 rapidly decays to zero as C1 charges. At the same time a positive step appears at the cathode, but this is conducted to ground by D4 and does not appear across R2. In a similar manner a negative step input causes a negative pulse to appear across R2. R1 and R2 may be connected in parallel without affecting the operation, since pulses can never arrive simultaneously via D1 and D2, and negative pulses will then appear at the output for a step of either polarity applied to the input. The amplitude of these pulses is determined by the amplitude of the step; their duration by the time constant R1C1 or R2C2.

The cathode-follower V13a is used, not only as such, but also as phase inverter to drive this pulse-forming device by inserting a suitable load resistor in its anode circuit in addition to the cathode load resistor. Since the cathode output is already loaded down by the transistor driven from it, these two load resistors have to be unequal in order to provide equal outputs to the diode-pairs.

Block T : C.R.T. Blanking-signal Generator. V14, D15, D16:

(a) Theoretical Considerations:

The output of Block S is a series of pulses, the amplitude of each indicating the rate-of-change of pitch at that particular instant. These can now be used to form a control signal which will suppress the C.R.T. spot when an excessive rate-of-change occurs, indicating an error in the pitch analogue. The required form of this control signal will now be discussed.

Consider a train of pitch period pulses as in Fig. 25a, containing gaps due to pulses which were missed as at t = 2 and also spurious pulses as at t = 8.75. These are the types of error which occur most frequently, and, apart from which, the pitch is assumed to remain constant at 180 Hz. Starting at t = 0, the pitch analogue is seen to remain constant at 180 Hz until t = 3, when it drops to the 90 Hz level since the pulse at t = 2 was omitted and hence the period from t = 1 to t = 3 measured. At t = 4 the analogue returns to the 180 Hz level and remains there until the occurrence of the spurious pulse at t = 8.75. The elapsed time from t = 8 to t = 8.75 is 0.75 x 1/180 second corresponding to a pitch of 4/3 x 180 = 240 Hz, which is correctly shown on the pitch analogue, but which should be suppressed in the final display. At t = 9, another normal pulse occurs and the time elapsed is only 0.25 x 1/180 second and the indicated pitch consequently 720 Hz, to which level the analogue rises. At t = 10 it returns to the 180 Hz level, where it remains.

The rate-of-change of pitch is shown in Fig. 25c, from which it is seen that all these sudden changes in the analogue produce pulses which exceed the indicated threshold level and will, therefore, actuate the blanking-signal generator.

Ideally this blanking signal should last only as long as the errors in the pitch analogue, but there seems to be little hope of reliably establishing the instant at which the erroneous signal terminates. It was reasoned, therefore, that a practical solution to the problem would be to generate a blanking pulse, which commences at the instant of the first trigger pulse and endures until a fixed time after the last. This margin after the last trigger is provided to ascertain whether another trigger is forthcoming or not, and has to be longer than 1½ times the duration of the longest pitch period for the following reason:

Take the case of a 90 Hz pulse train into which an extra pulse is injected at t = 4, as shown in Fig. 26a. In this special case the period measured at t = 4 precisely equals that measured at t = 5, namely 1/180 second corresponding to a pitch of 180 Hz, as shown in Fig. 26b. Fig. 26c shows a trigger pulse occurring at t = 4, but of course none occurs at t = 5 (an instant at which one of the genuine 90 Hz pulses occurs) since the pitch analogue stays constant at the 180 Hz level assumed by it at t = 4. Hence, if the blanking pulse lasts for less than 1.5 x 1/90 second, say for only 1.2 x 1/90 second as shown in Fig. 26d, the C.R.T. spot will reappear at t = 6.4, indicating a pitch of 180 Hz until t = 7 when the last trigger pulse occurs. If the spurious pulse is not located precisely half-way between t = 3 and t = 5, but at t = 4.1, say, the period from t = 3 to t = 4.1 is 1.1 x 1/180 second and that from t = 4.1 to t = 5 is 0.9 x 1/180 second. Hence the pitch rate-of-change detector will produce pulses both at t = 4.1 and at t = 5, but the pulse at t = 5 decreases in size as the spurious pulse approaches the precise centre between the pulses at t = 3 and t = 5. At some point near the central position, the difference in period will be insufficient to produce an error signal at t = 5. Hence the blanking pulse should be lengthened to endure for just more than 1.5 x 1/90 second after the occurrence of the last trigger pulse.

At higher pitch frequencies the above blanking period will suppress more cycles of the speech wave than strictly necessary since it remains constant regardless of pitch. However, this is of small consequence, since the speed at which speech sounds are produced is not a function of pitch. If a given sentence is therefore first produced by a bass voice and thereafter by a soprano voice <u>at the same speed</u>, and if the same errors occur and are suppressed in both cases, then the gaps in the resultant pitch displays will be of equal length and therefore equally tolerable.

(b) <u>Practical Details</u>: The blanking signal generator is required to produce a signal which will suppress the C.R.T. display. Suppression should start at the instant of the first trigger pulse produced by the pitch rate-of-change detector and should endure for a time just longer than 1.5 x 1/90 second (say 17 Msec.) after the last trigger. Only triggering pulses which exceed a preset threshold should produce an output from this device but, once triggered, the amplitude and duration of the blanking pulse should be independent of the amplitude of the triggering pulse.

It follows that a one-shot multivibrator which shapes all input triggers exceeding a certain threshold amplitude into short pulses of fixed duration and amplitude can be used to feed a timing circuit which will prolong the duration of all pulses for the required period of time. The action of this device is explained with reference to Fig. 27.

Negative-going trigger pulses A are applied to the grid of V1 via the diode D1 and are reproduced in inverted and amplified form at the anode of V1 as shown in C, since V1 is conducting during the time that the multivibrator is in its stable state. The grid of V2 is biassed far beyond cut-off and is coupled to the anode of V1 by a resistor. When a pulse of sufficient amplitude to cause conduction in V2 appears at its grid, the astable state is initiated. V2 anode current causes a negative-going pulse to appear at its anode, as shown in E, and this is coupled back to V1 grid by a small condenser which, in conjunction with the grid leak of V1, forms a timing circuit (RC = 1.8 Msec.). Fig. 27 B shows the grid voltage of V1 being carried far beyond cut-off, thereafter dropping back exponentially towards the grid voltage at which anode current in V1 once again starts flowing, causing a drop in V1 anode voltage,which leads to plate current cut-off in V2 and the conclusion of the astable state. Since the grid leak of V1 is returned to the + 150 volt line and the grid voltage swings from about -50 to about -4 volts during the astable state (-4 volts being the bias required for plate current cut-off when Va = 150 in an ECC 81 valve), discharge of the coupling condenser takes place over a relatively linear portion of the exponential curve and timing is reasonably accurate. The purpose of D1 is to disconnect V1 grid from the source of trigger pulses, once the astable state commences and so to avoid interaction between the two circuits.

One pulse of constant amplitude and duration now appears at the anode of V2 for each trigger of sufficient amplitude applied to the input, the threshold being set by the extent by which V2 grid is biassed beyond cut-off. This pulse is now further shaped in a circuit consisting of two diodes D2 and D3, a storage condenser of 0.04 MF and a discharge resistor of 2.2 megohms. This circuit forms a pulse which carries the grid of the C.R.T. into the region beyond cut-off of the electron beam for about 17 Msec. following each pulse of the multivibrator as shown in Fig. 27F. As in the case of the V1 grid circuit, the discharge is reasonably linear, since effectively only a small portion of the entire discharge curve is used. When the point F slightly exceeds the potential of the tap on the "brightness" control, the diodes D2 and D3 conduct preventing any further increase in potential. The 1K ohm resistor in series with the storage condenser allows the potential at F to drop initially at a rate as fast as the rise-time of the pulse at E, thus extinguishing the C.R.T. spot within a very short time. Since the rise-time of the radial deflection system is relatively long (see page 31) a negligible portion of the vertical stroke associated with an error in the pitch analogue is displayed. This is illustrated in Fig. 28 a, b and c.

It should be noted that it is not possible to dispense with the pulse lengthening system, designing the multivibrator to produce 17 Msec. pulses directly, since the action of a multivibrator is such that, once it is in the astable state, it is immune to external triggering pulses and hence it would not meet the requirement that the output pulse should endure for 17 Msec. after the last input pulse.

The width of the multivibrator pulses is not critical, but should be less than the smallest anticipated spacing between trigger pulses. This is unlikely to be less than 1/720 second because no signal components above 720 Hz enter the waveform shaping stages. Hence a pulse duration of about 1 Msec. is a reasonable design figure.

The function of the 1 MF decoupling condenser at the arm of the "brightness" control is to stabilize the mean bias voltage applied to the C.R.T. grid. Blanking pulses developed at F cannot change the mean charge on this condenser, since no D.C. path exists between it and the multivibrator.

The triggering threshold is set empirically by suitable choice of resistance values in the grid circuit of V2, using as criterion the fact that the fastest glides in pitch likely to be encountered, should not cause blanking whilst errors in the pitch analogue should be suppressed.

Block U : Cathode Ray Tube Display:

An outline of the display system finally adopted was given on page 11. . Some further details will now be described.

From several types of C.R.T. available the Ferranti type 12/03 YB was selected. This tube employs magnetic deflection and electrostatic focussing, has a controlled-afterglow type of screen and provides sufficient image brightness for viewing under conditions of subdued daylight.

Electrostatic focussing was preferred, since it merely called for a preset potentiometer across the 280 volt supply to the first anode of the tube to provide a focussing voltage on the third anode, whilst eliminating the need for a costly and energy consuming focussing coil assembly. By mounting the focussing potentiometer (in the form of two selected fixed resistors) on the holder of the tube base, the need for an additional slip ring for the third anode connection was eliminated.

A final anode potential of 12 KV is recommended by the makers of this tube and was supplied by an E.H.T. generator of the high-frequency oscillator type. This unit consists of a small power oscillator running at a frequency of about 150 KHz and feeding a transformer with a well-insulated secondary which produces the required high voltage. This is then rectified in a small high-voltage diode, the filament of which forms the positive terminal of the E.H.T. unit and is energized by a well-insulated winding on a separate transformer which is driven by its own power oscillator. By using separate circuits, it becomes possible to adjust the E.H.T. to any required value simply by changing the voltage applied to the relevant oscillator without affecting the filament voltage. The specifications of this unit are as follows:

Input: 300v at 70 Ma. max.; 6.3v at 0.9 amp.
Output: 1.5 to 15 KV at 300 microamps.
R.F. ripple on E.H.T.: Less than 1%.

Unit supplied by Messrs. Hivolt Ltd. of 91-93 Princedale Road, London, W.11, England.

Further circuit details are shown in the main circuit diagram, Fig. 7.

The C.R.T. final anode is connected to an internal conducting layer which, in conjunction with an external coating, forms a condenser which assists in the smoothing of the E.H.T. supply. The external coating is connected to the electrical ground via the ring gear which carries the tube.

The E.H.T. connection to the tube is a small recessed cap on the curved part of the glass behind the screen. A special spring-loaded connector which slides on the inside surface of a ring, serves as an E.H.T. slip ring, whilst a rotating bakelite disc with five concentric slip rings on its inner surface which slide over fixed carbon brushes, furnishes the necessary connections for filaments, cathode, control grid and first anode.

Since the neck of the C.R.T. is located close to the iron cored components of the power supply, it was necessary to provide magnetic screening around it. For the same reason, the deflecting coil assembly was enclosed in a magnetic screen in the form of a cup which extended some distance up the conical portion of the tube. This cup was made accurately coaxial with the neck of the tube and its angular position could be adjusted to ensure deflection of the electron beam into the required direction.

The C.R.T. face is coated with two distinct layers: that nearest the observer is a storage type phosphor, and the backing layer is a material capable of converting electron energy into blue, violet or ultra-violet light. It is principally the light from the backing layer (and not the electron beam directly) which activates the storage phosphor, and since daylight also contains these colours, it is important to shield the tube face from ambient light as far as possible in the interest of obtaining good contrast in the stored images.

The stored image is not immediately visible, but is released in the form of a yellow-green phosphorescence when the screen is flooded with infra-red light, this being a property of the zinc-sulphide phosphor used. The stored energy is limited, being derived from the original trace and under bright infra-red illumination is dissipated within 5 to 10 seconds. However, during this time the brightness of the stored image is sufficient for comfortable viewing under subdued daylight conditions, whilst at the normal speed of rotation of the C.R.T. the stored image has, for practical purposes, completely disappeared after one revolution.

Infra-red illumination was provided by a 15 watt lamp behind an infra-red filter which passes a wavelength of about 1 micron as recommended by the tube manufacturers. Lamp and filter were enclosed in a light shield, so positioned relative to the C.R.T. screen, that the viewing sector was both screened from ambient light and was suitably flooded by the infra-red light to produce fairly constant intensity of afterglow over its entire area. Since the images are recorded on the screen at the extreme right-hand side of the viewing aperture and are carried towards the left by rotation of the tube, the stored energy level decreases towards the left-hand side and hence the infra-red illumination has to be more intense on that side.

A switch was provided to stop the rotation in order that information stored on the screen can be viewed for several seconds before the patterns fade away. A second section on the same switch suppresses the electron beam, whilst the tube is stationary to prevent screen burning.

POWER SUPPLY:

General Considerations:

(a) Since the power supply is used in close proximity to the C.R.T., stray magnetic fields should be eliminated as far as possible. To this end a power transformer, using a ring core and symmetrical windings, is preferred. Choke input to the ripple filters should be avoided, since such a choke carries a large A.C. component with a resulting large leakage flux. High C filters are preferred.

(b) Reasonably good voltage regulation is required, since many stages obtain their power from the common supply and the current drain of some stages varies with the input signal. By using large smoothing condensers, good dynamic regulation and adequate ripple filtering are readily obtained.

(c) High efficiency A.C. to D.C. conversion obtained by using silicon rectifiers, ensures a minimum of heating, thus easing ventilation problems and allowing a smaller power transformer to be used. Their use also eliminates the need for rectifier filament supplies and ensures good voltage regulation.

(d) At the time this portion of the instrument was designed, silicon rectifiers having high P.I.V. ratings were both expensive and difficult to obtain. High-current types with moderate P.I.V. ratings were freely available, making the voltage-doubler rectifying system an economical proposition. At the same time, this leads to very modest voltage requirements at the power transformer secondary.

SPECIFICATION:

The following power requirements should be met by the power supply under conditions of continuous duty:

+ 280 volts at 200 Ma.	:	56 w)		
- 140 " " 50 "	:	7 w)	Total =	64 watts D.C.
- 16 " " 60 "	:	1 w)		

This includes currents drawn by the + 150 volt and - 105 volt regulated supplies which are derived from the + 280 volt and - 140 volt lines respectively.

Valve filament current totals 7 amps and a 6.4 volt winding was provided to allow for wiring losses. Filament power is therefore 45 watts.

Allowing for losses in the A.C. to D.C. conversion and smoothing processes, and in the power transformer itself, leads to the choice of a core which can accommodate 150 watt windings.

The required voltage and current ratings for the H.T. secondary of the transformer are determined from published semiconductor diode data and are shown in Fig. A7 Two identical bobbins carrying these windings were made and placed on a ring core. Corresponding windings on the two bobbins were series connected, thus ensuring that the same currents flow in both windings and that the flux in the two legs of the core balances. This procedure ensures a minimum of stray flux. In addition, a flux density of only 75% of that which is considered normal for the core, was used. Finally the transformer was enclosed in a box forming a magnetic screen.

Waveshapes present at 9 points in the circuit when the word "soot" is spoken, are shown in Fig. 28c.

- - - oOo - - -

CHAPTER 9.

MECHANICAL DETAILS.

CATHODE RAY TUBE SUPPORT:

The 12" diameter cathode ray tube is held in a specially cast brass ring gear assembly, the inner surface of which was machined to match that of the C.R.T. over the curved portion of the glass near the screen of the tube. This assembly supports the bulk of the weight of the tube, but another flexible point of support was provided at the end of the neck, this support at the same time serving as a slip ring assembly for the required five electrical connections to the tube base.

Front and rear supporting rings were both grooved to take the outer races of ball bearings forming the rotating points of support for the rings. Four ball bearings support the front ring gear and three the back slip ring assembly, and each is mounted on a pillar, the other end of which clamps down on the supporting framework. The threaded studs of these pillars are not concentric with the centre of the ball races and hence by turning the pillar, after loosening its lock nut, the position of the ball bearings relative to the framework may be slightly changed. Hence the pressure between each ball bearing and the ring supported by it may be accurately set, and alignment of front and back rings may be carried out.

Since perfection in the contour and symmetry of the C.R.T.

could not be expected, the ring gear inner contour was machined with about $\frac{1}{8}$" clearance and the tube was protected from hard contact with the brass by nesting it inside a knitted woollen jacket which ensured adequate radial support. In addition, a soft plastic tube was laced to the front inside of the ring to serve as a buffer which prevents longitudinal movement forwards. See Figs. 29 and 30.

The frictional force between the tube and the ring is sufficient to transmit the required torque for the rotation of the tube. The required electrical contact between the conductive coating on the tube and the ring gear is provided by an earth strap made of flexible wire which is pressed down against the glass by adhesive tape. This is visible in Figs. 31 and 33.

Further details of the ring gear which can be separated into two parts when the circumferential screws (visible in Fig. 29) are removed, are presented in Fig. 36.

REAR SUPPORT AND SLIP RING ASSEMBLY:

To relieve strain on the tube neck, the back support is made flexible in the following manner:

As shown in Fig. 32, the tube base is plugged into a mating socket (B12A duodecal) which is centrally held in a large hole in the synthene disc of the slip ring assembly by three spiral springs stretched radially between three points on the circumference of the socket and three anchoring points on the disc. In addition to supporting the neck of the tube, these springs also have to transmit torque between the tube and the disc to overcome the friction of the three ball bearings surrounding the disc and that of the electrical brushes. To suppress a tendency towards mechanical oscillation in the angle by which the disc lags behind the tube neck, the springs were brought into frictional contact with rubber blocks attached to the disc as shown in Fig. 32. The other slack leads seen in this photograph are the electrical connections to the slip rings. These are soldered to the countersunk heads of the screws which hold down the five concentric brass slip rings. The latter are out of view on the opposite side of the synthene disc. The brushgear is seen in Fig. 30.

DRIVE SYSTEM:

There are 340 teeth machined onto the ring gear which meshes with a 17 toothed pinion driven by a governor type of gramophone motor as seen in Figs. 30 to 33. The motor speed is variable between about 20 and 80 R.P.M., making available speeds of rotation of the C.R.T. between 1 and 4 R.P.M. At the radius of the centre of the viewing window ($4\frac{3}{4}$") this results in a screen speed range of $\frac{1}{2}$ to 2 inches per second. Practical experience indicates that the middle of this range is optimum. Fig. 36 shows a section through the ring gear.

It was found imperative to provide flexibility in the coupling between the motor and the ring gear. Without this, the high mechanical inertia of both driving and driven members allowed tooth ripple inherent in this type of spur gear drive to set up forces in the structure which supports the ring gear and motor. These forces oscillated at the frequency of the tooth ripple and at certain speeds of rotation, mechanical resonances in the structure allowed the resulting vibrations to assume alarming proportions accompanied by much unwanted noise and random fluctuations in the speed of the ring gear. Figs 30 and 33 show this flexible shaft in position between motor and pinion, and Fig. 37 shows details of its construction.

FRAMEWORK:

This is $14\frac{3}{4}$" wide, $14\frac{3}{4}$" high and $19\frac{1}{2}$" deep and made up of $\frac{3}{4}$" x $\frac{3}{4}$" x $\frac{1}{8}$" angle iron welded at the corners, extra strength being imparted by diagonal pieces across the front and bottom frames. These pieces are visible in Figs. 29 to 31. Two $\frac{3}{4}$" x $\frac{1}{8}$" flat vertical struts support the slip ring assembly, itself assembled on a circular metal disc of 7" diameter and $\frac{1}{8}$" thickness. The driving motor and vertical supporting structure for the deflection coils are screwed to a tray running across the bottom framework, as seen in Figs. 30 to 33.

Fig. 34 shows the two chassis on which the electronic components are mounted. These are housed in the spaces above and below the neck of the C.R.T. and are bolted to the framework, as shown in Fig. 35.

To remove or insert the C.R.T., the instrument is tilted over to stand on the rear frame and the twelve screws which clamp the two halves of the ring gear together are removed. This allows the top half to be lifted off, exposing the tube which is then free to be lifted up vertically, whilst the tube pins will be extracted from the socket. To aid reassembly, there are two blank holes into which locating pins are inserted to hold the two halves in position, whilst the twelve screws are replaced.

E.H.T. SLIP RING AND BRUSH:

The E.H.T. slip ring consists of a cast brass ring shown in Fig. 36. It is held in position, as shown in Figs. 29 to 33, by three conical insulators, the bottom one being screwed to the central tray and the top two to the horizontal frame members. The ring is located precisely over the recessed E.H.T. connector which is situated in the conical part of the C.R.T. bulb, and electrical connection between the two is provided by a brass saddle which slides along the inner surface of the ring and which is connected by means of a bowed spring wire to a plug which fits into the recessed cap in the C.R.T. To prevent the formation of a corona discharge, sharp edges are avoided on all parts carrying E.H.T. For the same reason the spring wire is covered in an insulating tube. Details of the saddle and plug are shown in Fig. 38.

FRONT COVER:

This serves as a mask over the face of the C.R.T., of which only a 80° sector (about 6" long) by $2\frac{1}{2}$" high, is visible in the viewing window at the top, as shown in Fig. 39. The DM 70 unvoiced sound indicator is housed in a U-shaped member sweated horizontally across the centre of the cover with the purpose of stiffening the metal. Thus the indicator is visible in a slot immediately below the viewing window where it may be seen by peripheral vision, whilst the user is concentrating on the pitch patterns. Over the top of the viewing window is a lamp house for the infra-red light source - a 15 watt 230 volt "Pygmy" lamp enclosed in a metal tube, the inside of which has been given a bright finish and into the lower side of which an oblong slot is cut. The slot is covered with an infra-red filter and is suitably positioned relative to the C.R.T. face to provide the required distribution of infra-red light across the area of the viewing aperture. The latter is covered with a sheet of clear celluloid serving as a protection for the C.R.T.

Below the central stiffening member is a 7" x $8\frac{1}{2}$" mirror hinged about its top edge and supported by a curved rod which passes through a hole at the bottom of the front cover. A leaf spring is in frictional contact with this rod just behind the front panel, thus ensuring that the mirror will remain standing at any desired angle.

Controls and sockets required by the operator are brought out on the front panel, as shown in Fig. 40. In clockwise order, starting from the top left-hand corner, these are:

Pitch/volume selector, earphones volume control, main on-off switch, earphones outlet, microphone input and the switch for stopping the C.R.T. rotation.

In addition the following preset controls are available for screw driver adjustment: C.R.T. brightness and spot position on left-hand side of back cover and gain control on the right.

The back cover can be taken off after eleven retaining screws and the carrying handles have been removed. To facilitate this operation, two extracting bolts are provided at the back. For extraction these should be screwed in a clockwise direction.

To aid in noise reduction, the instrument stands on four rubber feet, bolted to the diagonal stiffening pieces in the base frame.

- - - oOo - - -

CHAPTER 10.

SOME PRACTICAL RESULTS AND CONCLUSIONS.

At the time of writing, the instrument has been in use at the School for the Deaf, Worcester, for more than a year. Children of all grades have had the opportunity of becoming acquainted with it, including those of nursery school age. Initially, teachers were shown how to use the instrument and thereafter they were left to discover new avenues of approach and new applications for themselves. They were encouraged to describe their experiences in a note book which could be consulted by the others for their mutual benefit.

After several months of use, the continuous time base was changed to the manually triggered single-sweep time base (see pg. 18) and the teachers were requested to report on the children's reactions.

TEACHERS' REPORTS:

(a) The children are very enthusiastic about the instrument, even after the first novelty has worn off, since they are able to show continuous progress.

(b) The simultaneous use of the senses of touch and sight proved to be of value, since it results in a keener awareness of the existence of sound.

(c) By using the sense of touch to follow the sound patterns of another speaker, the children are enabled to devote their full visual attention to patterns on the screen. Consequently group work can be done which would otherwise be difficult since it is impossible to watch the speaker's lips and the screen simultaneously. See Fig. 41.

(d) Group work encourages a spirit of competition which results in rapid progress, since the children learn from each other's mistakes.

(e) Amplitude pattern display is very useful at all times, but especially with beginners who are generally only able to grasp the meaning of pitch after a prolonged period of practice.

(f) Rapid switch-over from amplitude to pitch display of the same utterance enables the two patterns to be compared. Thus it becomes possible to demonstrate the interdependence of stress and pitch patterns; this serves as an introduction to the concept of pitch.

(g) Used under subdued daylight conditions with the instrument so placed that the screen is as dark as possible and the teacher's face well visible, the persistence of the screen has proved to be adequate. Patterns remain visible for about ten seconds when the rotation is stopped and teachers report that it is seldom considered desirable to have more time available for studying the patterns.

(h) In fact, the stop switch is not used frequently, since at the speed of rotation which was finally adopted, namely about 1" per second at the centre of the viewing aperture, there is sufficient time to study a pattern as it traverses the aperture.

(i) If finer resolution of the C.R.T. spot can be achieved,* the time base can be slowed down further and patterns compressed, allowing more information to be accommodated in the aperture. The teachers suggested that this would be useful.

(j) It goes without saying that the persistent images are found absolutely essential as a means of portraying pitch or amplitude information. This could readily be proved by switching off the infra-red illumination. This reduces persistence to only one to two seconds. As a further test, spot intensity is reduced until the only luminescence which remains visible is that of the short persistence blue spot, the rapid radial movements of which are practically meaningless to the observer. Even if, in this case, the single-sweep time base is used, the eye finds it difficult to memorise the two-dimensional spot movements.

(k) The unvoiced-sound indicator is especially useful for showing the <u>s</u> sound which most deaf children find difficult to use correctly. It also shows a strong <u>t</u>.

(l) There has never been any question about the equivalence of the visual patterns to aural pitch patterns showing that the error suppression does not introduce any too obvious artifacts.

<u>COMPARISON BETWEEN CONTINUOUS AND SINGLE-SWEEP TIME BASES:</u>

Depression of a thumb switch in the handle of the microphone caused the C.R.T. spot to start its sweep at the left-

---

* A trace of stray magnetic field which still exists in spite of the precautions taken to prevent it, limits the resolution in the present instrument.

hand edge of the viewing window. It then travelled across the screen at a speed equivalent to that at which the screen moves past the window in the case of the continuous time base. It follows that the patterns were visible for the same length of time in the two cases.

The basic difference between the two systems lies, of course, in the fact that in the case of the continuous time base the child may start speaking at any instant and continue for as long as he desires, whilst in the case of the single-sweep time base he has to co-ordinate the commencement of his speech with the action of depressing a switch and he can only continue talking for six seconds before the display disappears behind the mask at the right-hand side.

A second, less important, difference lies in the curved vs. the linear shape of the spot movement.

After a considerable period of use, the following views were expressed by the staff who used the instrument:

(a) Co-ordination between the actions of pressing the switch and starting to talk is only achieved after a period of practice, very young children and retarded children find this quite difficult.

(b) It was found that most children could readily adapt themselves to either the curved or the linear base lines inherent in the two types of display. Apparently the lower edge of the viewing aperture guides the eye in its assessment of the radial position of the trace.

(c) An important objection to the single-sweep time base is its time limitation. When a child is slow to react or when an exercise requires a long sustained series of utterances (e.g. to teach breath control) six seconds is insufficient time.

(d) On the other hand, more than ten seconds duration is seldom required. If a longer viewing window could be arranged, or alternatively, finer C.R.T. resolution and a time base of longer duration, the objection raised in (c) above, would not be so important.

(e) Advanced children could adapt themselves to either type of display with equal facility and found both acceptable.

(f) It was felt that the need to press a button formed a distraction which, in the case of beginners, retarded progress.

CONCLUSIONS:

It may be concluded, therefore, that finer C.R.T. resolution and a slower time base speed to give, say, 10 to 12 seconds transit time across the screen would enable the same size of display to be used.

The persistence time of the screen phosphor is satisfactory, but brightness of patterns should be no less than at present.

Both types of display are acceptable for advanced work, but if the instrument is to be of maximum value to children of all grades, the continuous time base is preferred.

In view of the remarks contained in (b) above, provision of a graticule, finely ruled with concentric arcs in the case of the continuous time base or with parallel straight lines in the case of the single-sweep time base, placed over the viewing window may be an advantage.

The method of pitch analysis appears to be satisfactory and the pitch analogue sufficiently accurate for the present application.

The need to select a suitable pitch range manually, causes no difficulty, since the only effect of an incorrect selection is to produce more blank spots in the patterns. The action of the error detection and blanking circuit is satisfactory under practical conditions of use.

APPLICATION NOTES:

The application of the instrument to the training of deaf scholars was outlined in Reference (1), pg. 1073. Even at that stage it was clear that a deaf child had great difficulty in grasping the meaning of pitch variations, and the training program in use at the time was designed to prepare the child for an ultimate understanding of this concept.

In the new instrument the facility for displaying amplitude patterns was included and this made possible the development of new exercises which will be described briefly.

A. Correction of Individual Words:

1. Deaf children are inclined to draw out voiced sounds e.g. "a caaat" instead of "a cat". It can be demonstrated that the a starts and stops abruptly and is of short duration.

2. Extra syllables are often injected, e.g. "tooul", "su-pring" and "hat-us" instead of "tool," "spring", and "hats", or added at the end of a word, e.g. "cat-u", "sat-u" etc. These show on the amplitude pattern and can be corrected. The s and t sounds are also clearly seen on the unvoiced sound indicator.

3. Voiced consonants are extended unduly, e.g. "falll", "moonnnn", "farmmmm". Usually the amplitude pattern shows the transition between the vowel and the voiced consonant and the respective lengths of these can therefore be seen.

B. Breath Control:

Unlike children with normal hearing who learn at an early age to breathe only at natural pauses between words, deaf children have to be taught how to control the flow of their breath in forming words and rules have to be laid down for them to follow. Unless this is done, they are quite capable of taking a breath in the middle of a word - to quote a common tendency. This, of course, destroys intelligibility. In teaching breath control, the following exercises have been found useful:

1. Maintaining a constant amplitude for as long as possible until the breath runs out on a vowel like "ooo..." or "eee...", or a voiced consonant like "mmm...", "lll..." etc.

2. Getting short bursts of sound all evenly spaced (in time) and all of the same amplitude, e.g. "b b b...", "d d d ...", "th th th...", etc.

3. As for 2., but using longer sounds, e.g. "baa baa baa...", "dee dee dee ...." etc.

4. Combinations of 2. and 3., e.g. "b b baa, b b baa..." or "b baab, b baab ..." etc.

5. Forming a series of breath consonants all of equal amplitude, e.g. "p p p ..." or "t t t ..." etc.

Later the lessons so learnt are extended to forming words consisting of short and long vowels, e.g. "the c<u>a</u>t, the m<u>oo</u>n".

C. <u>Teaching of Correct Rhythm in Speech</u>:

This is the next logical step after breath control exercises. Instead of merely using sounds, words and sentences are formed having the same rhythmic patterns.

1. Words with short vowels, e.g. "a cat", "a pot" etc.

2. Long vowels, e.g. "blue moon", "green peas", etc.

3. Longer words with syllables of unequal length, e.g. "a table", "a tomato", etc.

4. Forming short sentences, e.g. "It is hot."

D. <u>Teaching Accent</u>:

This is closely related to rhythm and the same procedure is followed. The child can be shown that the timing, as well as the size of the individual parts of the pattern, have to be correct. To illustrate: it can be considered normal practice to accentuate the underlined syllables in the following expressions. "a <u>cat</u>", "a <u>ta</u>ble", "to<u>ma</u>to". The deaf child may stress the incorrect syllable, e.g. "a ta<u>ble</u>", "a <u>to</u>mato", which has a detrimental effect on the intelligibility of his speech. Using the instrument, the correct and the incorrect patterns may now be demonstrated to him and he can practice and immediately see the results of his own vocal efforts and compare these with the standard as set by the teacher.

When the child has advanced to this stage, he is usually ready to start work on pitch.

E. <u>Developing the Concept of Pitch</u>:

It was pointed out in Reference (1), pg. 1065, that the concept of pitch is possibly as abstract to a totally deaf child as the concept of colour is to one born blind. To such children the existence of these qualities can only be explained in terms of some concept or some sensation with which they are already familiar. In the case of pitch, the child's attention can be drawn to the fact that different pitches correspond to different degrees of tension in the muscles of his throat. Using the pitch indicator, the results of his muscular exertions now become visible and he learns to correlate the two. Given sufficient exercise in front of the machine, he is later able to carry over his newly gained voice control into his everyday speech.

Some useful suggestions were made by the teachers in connection with the introduction of the concept of pitch to a beginner.

1. Some words are easier to pronounce if the correct pitch variations are used, e.g. "water" and "table". Most deaf children are able to get some change of

pitch between the two syllables of such words. Using the pitch display, it can be demonstrated to them that they are already able to get a pitch variation and a connection between the tactile sensation in the throat and the visible pitch pattern is formed by the child.

2. Once the child has formed this association, he is ready for the following exercises:

   (a) Getting two levels of pitch on a sustained vowel, e.g.

   oo oo oo
   oo oo

   (b) Imitating animal noises:

   Cow: Mmoooooooooooooo

   Donkey: Hee Haw

   or other gross sounds, e.g. bell: ding dong

Using simple sounds as indicated above, allows the child to devote his full attention to developing his pitch control. Since the child is by now familiar with the concepts of low and high pitch, as portrayed by the instrument, the teacher can indicate to the child what changes are required by suitable hand or head movements.

More advanced work consists of applying the above basic experience to phrases and sentences.

Fig. 42 illustrates some typical amplitude and pitch patterns obtained by saying the sentence "his cat is brown" in four different ways in which the stress is in turn placed on each of the four words, thus conveying different meanings to this simple statement. The importance of correct stress and pitch control is clearly indicated.

- - - - - oOo - - - - -

APPENDIX.

FILTER DESIGN:

The design of the active R-C type of filter which follows, is fully described in Reference 17, where it is shown that if a damping factor Df of more than 0.5 is used and if the cathode follower which constitutes the active part of the network has a gain which is close to unity then in Fig. A4:

$$2\,Df = C2/C1 + 1 = R1/R2 + 1 \quad \dots\dots\dots\dots\dots\dots \quad (1)$$

$$\text{and} \quad R1C1 = R2C2 = 1/Wo = \text{corner angular freq.} \quad \dots\dots \quad (2)$$

This is the same for both high-pass and low-pass filters. A satisfactory transfer function shape results if Df lies between 0.6 and 0.8.

Substituting these values in (1) produces C2/C1 values between 0.2 and 0.6.

For convenience C2/C1 was made 0.25 in all cases.
Hence C1 = 4C2 and R2 = 4R1 .... (3)

This choice of parameters also results in a transfer function which is little affected by changes in valve characteristics. The attenuation beyond the knee frequency is 12 dB per octave. By cascading several identical stages, any multiple of 12 dB per octave may be obtained.

Since 24 dB per octave is required, the two sections of an ECC81 valve were used in all filters. In all cases Rk = 47 K ohms, B + = +150 volts and B- = -105 volts.

HIGH-PASS FILTER: (See pg. 21)

$$fo = 4000\ \text{Hz}, \quad Wo = 2\pi \times 4000 = 2.52 \times 10^4.$$

Choosing C1 = 1000 pf (for convenience) then from (2):

$$R1 = 1/Wo.C1 = 1/2.52 \times 10^4 \times 10^{-9} \quad \therefore \quad R1 = 40\ \text{K. ohms}.$$

Hence from (3) R2 = 160 K. ohms and C2 = 250 pf.

- - - oOo - - -

BAND-PASS FILTER: (See pg. 22) 90 to 720 Hz.

This consists of a high-pass filter, fo = 90 Hz, in cascade with a low-pass filter, fo = 720 Hz.

H.P. FILTER: $Wo = 2\pi \times 90 = 565.$

For convenience make C1 = 0.02 MF = 2×10^{-8} F, then:

$$R1 = 1/Wo.C1 = 1/565 \times 2 \times 10^{-8} \quad \therefore \quad R1 = 88\ \text{K. ohms}.$$

Hence R2 = 350 K. ohms and C2 = 0.005 MF.

L.P. FILTER: $Wo = 2\pi \times 720 = 4520.$

For convenience make C1 = 0.002 MF = 2×10^{-9} F, then:

$$R1 = 1/Wo.C1 = 1/4520 \times 2 \times 10^{-9} \quad \therefore \quad R1 = 110\ \text{K. ohms}.$$

Hence R2 = 440 K. ohms and C2 = 500 pf.

Two each of these H.P. and L.P. stages were cascaded as shown in Fig. A5.

DISCHARGE OF CONDENSER: (See pg. 27)

If a condenser C is charged to a potential V and then allowed to discharge into a resistor R, the potential E left over the condenser at time t will be:

$$E = Ve^{-t/RC}$$

If t/RC is small compared with unity then:

$$E = V\,(1 - t/RC) \quad \text{or} \quad E/V = (1 - t/RC)$$

Hence $t/RC = (V-E)/V$ or $C = tV/(V-E)R$ .... (4)

On page 27: $(V-E)/V = 2\%$. Therefore $V/(V-E) = 50$
also: $t = 1/90$ and $R = 15 \times 10^6$

Hence, from (4): $C = 1/90 \times 50/15 \times 10^6 = 0.037\,\mu F$ (C4)

--- oOo ---

TRANSFER OF CHARGE BETWEEN CONDENSERS: (See pg. 27)

Consider the case of a large condenser, C, charged to a potential V and then placed in parallel with a small condenser, K. Thereafter K is isolated and C recharged to the same potential. Next, K is again paralleled with C and thereafter isolated. It is required that the second increment in the potential across K should be less than V/20. Determine:

(a) the minimum value of C/K;
(b) the value of the third increment in potential across K if the above process is repeated once more.

(a) The charge in C is VC coulombs. Initial charge in K is zero. When the condensers are paralleled, the total charge remains constant at VC, whilst the capacity becomes $(C + K)$. Therefore the potential drops to $VC/(C + K)$. When isolated, K retains this potential which corresponds to a charge of $KVC/(C+K)$ in K.

The charge in C is now restored to VC and C is again paralleled with K. Total charge is now:

$VC + KVC/(C + K)$, or, simplified: $VC.\left[1 + K/(C + K)\right]$

Hence the potential drops to: $VC.\left[1 + K/(C + K)\right] /(C + K)$

Increment in potential across K is therefore:

$$VC.\left[1 + K/(C + K)\right] /(C + K) - VC/(C + K)$$

This simplifies to:

$VC\left[K/(C + K)\right] /(C + K) = VCK/(C + K)^2$ .... (5)

Substituting the given value of this increment (V/20) in (5) leads to:

$C^2 - 18CK + K^2 = 0$, a quadratic in C, the solutions of which are:

$C = 17.94\,K$ or $0.055\,K$.

Hence $C/K = 17.94$ (the other solution being trivial.)

In practice, $C/K = 25$ was used ($C = 1$MF, $K = 0.04$ MF) and substituting these values in (5) above gives:

$$\text{Increment} = V \times 1 \times 0.04/(1 + 0.04)^2 = 0.037 \text{ V}.$$

The increment is therefore 3.7% of V, which is satisfactory.

(b) After the second increment, the charge on K was:

$VCK.\left[1 + K/(C+K)\right]/(C + K)$. If K is now again placed in parallel with C, which had been recharged to VC coulombs, the total charge becomes:

$$VC + VCK\left[1 + K/(C + K)\right]/(C + K) = VC\left\{1 + K\left[1 + K/(C+K)\right]/(C+K)\right\}$$

and the potential drops to: $VC\left\{1 + K\left[1 + K/(C + K)\right]/(C + K)\right\}/(C+K)$

The third increment in potential across K is then:

$$VC\left\{1 + K\left[1 + K/(C + K)\right]/(C + K)\right\}/(C + K) - VC\left[1 + K/(C+K)\right]/(C + K)$$

Which simplifies to $VCK^2/(C + K)^3$ .... (6)

Substituting $C = 1$ MF and $K = 0.04$ MF, the third increment in potential across K becomes:

$$V \times 1 \times (0.04)^2/(1.04)^3 = 0.0014 \text{ V or } 1/7\% \text{ of V}.$$

--- oOo ---

DESIGN OF PULSE TRANSFORMER: (See pp. 29 - 30).

When the original transformer design was performed, no suitable Zener diode was available for use in the secondary circuit of the transformer. It was experimentally determined that one plate of a selenium rectifier had suitable properties, but the "knee" voltage fell at about 50 volts (and in subsequent use this rose to about 80 volts, so that this unit had to be discarded). The "knee" voltage of the two OAZ 207 Zeners currently in use is only 17 volts. Because of this and other unknown factors, such as the width of the pulses which the transformer has to handle, the design had to be very conservative. Now that operating conditions are known with more precision, it would be possible to reduce the size of the transformer appreciably. The original design will be presented, since the principles involved remain the same.

ESTIMATING TRANSFORMER VOLTAGES:

Secondary: Assume a current of 100 Ma to allow a margin of safety over the estimated 50 Ma, an effective diode resistance of 200 ohms and a transformer series resistance of 100 ohms. Total voltage drop over circuit resistance is then 30 volts. The Zener voltage was assumed to be 50 volts and therefore the pulse amplitude has to be 80 volts.

Primary: This forms the load circuit of the pentode section of an ECL 80 valve operating with Va = 280 volts and Vg2 = 150 volts. At zero bias the anode current does not exceed 40 Ma. under these conditions and therefore, to get a 100 Ma pulse current in the secondary of the transformer, the turns ratio: primary/secondary has to be 100 /40 = 2.5. Hence, when 80 volts appear across the secondary, the primary voltage will be 200. From published data for this valve type, these conditions appear to be very reasonable.

Windings: To determine the required number of turns, the flux density in the core has to be considered. This is maximum at the end of the flat top of the pulse and a safe value which may be assumed for normal grade transformer steel is 50,000 lines per square inch. A maximum pulse duration of 0.5 milli-seconds was assumed. For the duration of a flat-topped pulse, the voltages across the windings are constant - 200 volts in the case of the primary. Consequently, the rate-of-change of flux linkages remains constant. Let flux linkage = G, then:

$$dG/dt \times 10^{-8} = E = 200 \text{ volts}$$

Therefore $dG/dt = 2 \times 10^{10}$ or $G = 2 \times 10^{10} \times \int_0^t dt$

When $t = 0$, $G = 0$.

When $t = 5 \times 10^{-4}$, $G = 2 \times 10^{10} \times 5 \times 10^{-4} = 10^7$

But G = nAB where n = number of turns
A = area of turn
B = flux density (5×10^4 lines/sq.inch)

Hence $nA = G/B = 10^7/5 \times 10^4 = 200$

A core of 9/16" x 9/16" was available. To allow for stacking, an effective core area (A) of 0.28 sq. inch can be assumed.

It follows that n = 200/0.28 = 714 turns (primary). Secondary turns therefore = 714/2.5 = 286.

To provide for unknown factors, 1000 turns were placed on the primary and three taps were provided, as shown in Fig. A6, to enable the turns ratio to be altered as required.

In the manufacture of the windings, the primary was first put onto the bobbin. This was then covered with a layer of insulation over which was placed an electro-static screen consisting of $1\frac{1}{4}$ turns of brass sheet as wide as the transformer window. The overlapping ends were insulated from each other and the screen was connected to the transformer core clamps and hence also to the chassis.

Over the screen a layer of insulation was placed and the available winding width was then divided in two, each half being filled with 143 turns of wire. In order to ensure equality of winding-to-ground capacity, one half of the secondary was wound in the opposite direction in order that the windings would be series-aiding if the two starting points were bridged.

The wire gauge used for all windings is 34 S.W.G., which fills the available space and results in an effective series resistance of only 25 ohms across the secondary windings.

LEAKAGE INDUCTANCE:

With the primary short-circuited and the secondaries series connected, the secondary impedance was measured at frequencies of 100 to 20,000 Hz. Knowing the effective D.C. resistance at the secondary, the leakage reactance and hence the leakage inductance could be found. This is 1.4 mH. It is wise to consider the effect of this on the current build-up in the secondary load circuit. The resistive part of the latter is now 200 ohms (diodes) plus 25 ohms (transformer). The time constant of a circuit consisting of inductance and resistance is L/R, which in this case is $1.4 \times 10^{-3}/225 = 6.2$ microseconds.

Since this time constant is negligible compared with the pulse duration, leakage inductance may be ignored.

- - - - - oOo - - - - -

V.T.V.M. MEASUREMENTS OF VALVE PIN VOLTAGES (ZERO SIGNAL CONDITIONS).

PIN NUMBERS.

Valve No.	Type	1	2	3	4	5	6	7	8	9	H.T. Supplies.
1	ECC83	120	0.5	0	0	0	220	–	30	6.3 AC	270
2	ECL82	0	21	0	6.3 AC	0	270	260	1.5	160	275
3	ECC81	220	0	3.2	6.3 AC	6.3 AC	150	0	2.6	0	280, 150, -15
4	ECC81	150	2.2	4.5	0	0	150	0	2.3	6.3 AC	150, -105
5	ECC81	150	2.0	4.1	6.3 AC	6.3 AC	150	0	2.1	0	150, -105
6	ECC81	200	0	1.3	0	0	200	0	1.3	6.3 AC	270
7	ECC83	140	0	2.3	0	0	16	0	0	6.3 AC	140
8	ECF82	240	16	150	6.3 AC	0	240	57	57	55	260, -105
9	EA50	0	280	6.3 AC	280	–	–	–	–	–	
10	ECL80	84	1.2	0	6.3 AC	0	280	0	150	-35	280, 150, -105
11	6 AL5	0	0	6.3 AC	0	0	–	0	–	–	–
12	6 AL5	0	0	6.3 AC	0	0	–	0	–	–	–
13	ECL82	0	-105	-135	0	6.3 AC	0	0	2.3	240	280, -135, -105
14	ECC81	23	0	0	0	0	150	-19	0	6.3 AC	150, -105
15	ECC82	110	0	3.4	0	0	110	0	3.4	6.3 AC	280
16	ECC81	150	0	2.0	0	0	150	0	2.0	6.3 AC	150, -105
17	12/03/YB	0	-48	–	–	–	280	210 (Pin 10)	2.6 (Pin 11)	6.3 AC (Pin 12)	280
18	DM70	-16	–	–	1.3 AC	0	–	–	85		

TYPICAL SIGNAL LEVELS.

The following table shows typical signal levels throughout the circuit. Sinusoidal input signals at 100, 200 and 500 Hz were applied to the microphone socket and a calibrated oscilloscope was used to perform the measurements. All results are given in peak-to-peak volts.

Point of Measurement:	Frequency (Hz) 100	200	500
Mike input	0.10	0.10	0.10
Ph. inv. grid	5.0	5.0	5.0
V3a plate	90	90	90
A	50	75	70
B	1.5	3.8	3.7
C	1.5	2.6	2.0
D	24	36	29
E	2.0	4.0	3.2
F	1.6	3.1	1.4
G	28	36	24
H	3.6	4.0	3.5
I	3.3	3.0	1.1
J	50	50	50
K	27	28	28
L	40	40	40
M	120	110	100
N	37	38	37
O	33	34	31
P	8.0	6.0	3.0

- - - oOo - - -

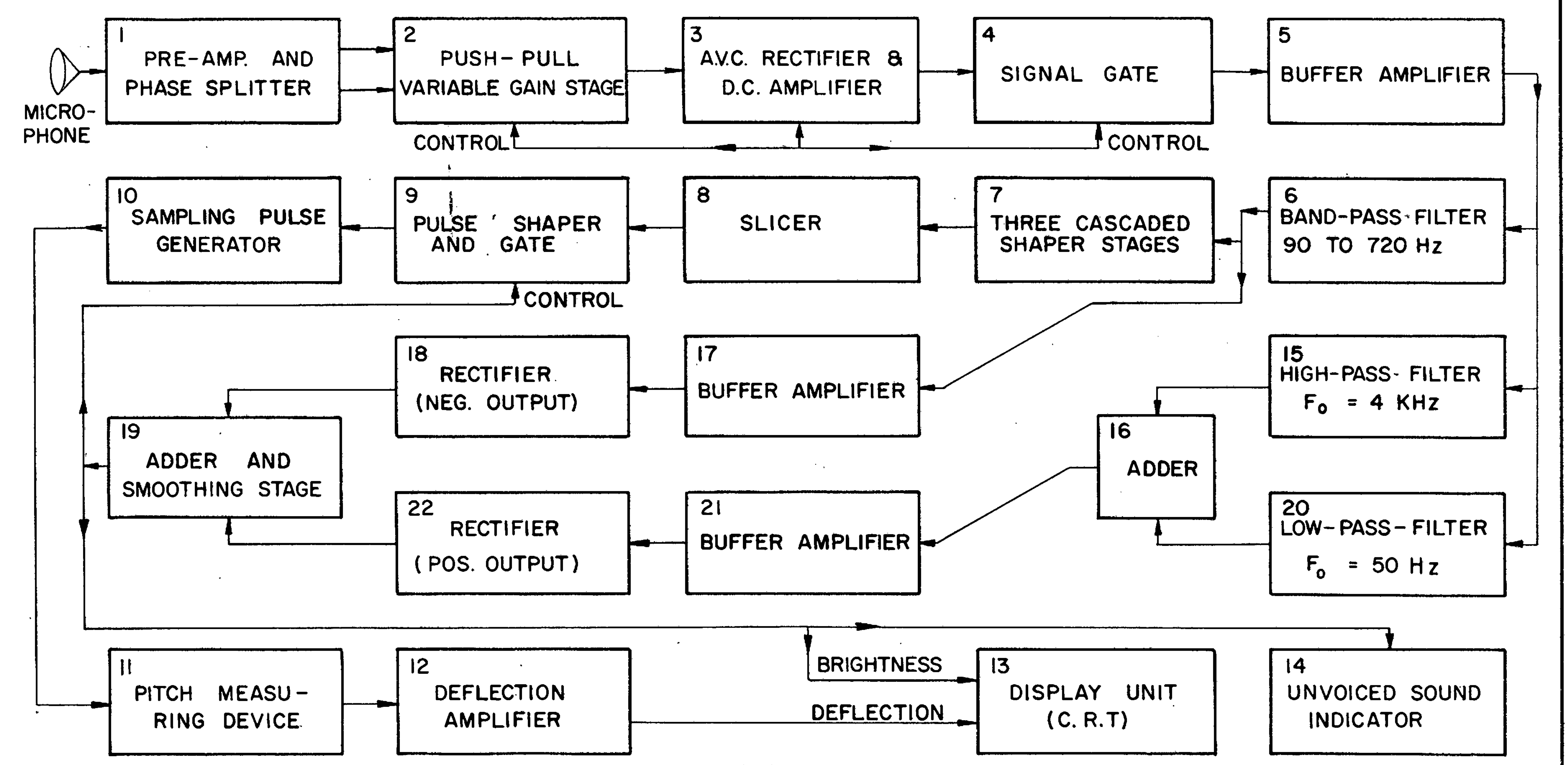
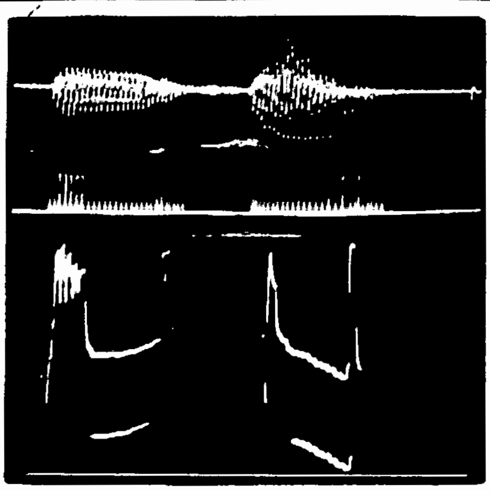
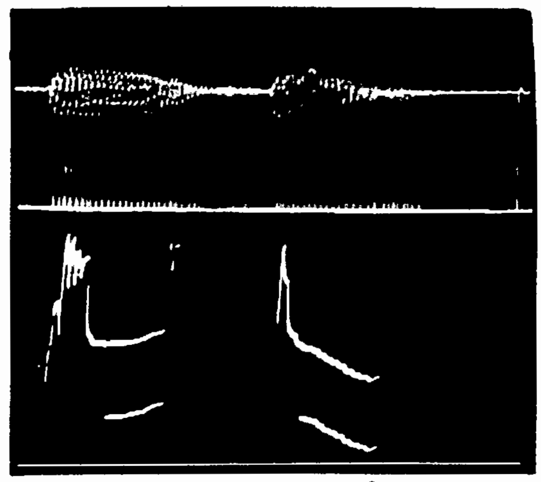


FIG. 1 . *Block diagram of preliminary design.* *(Pp 10-12)*



a
b
c
d



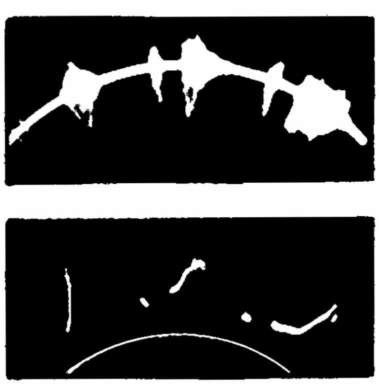
S E A SH O RE

FIG.2 WAVESHAPES PRODUCED BY THE ABOVE WORDS.
a: MICROPHONE OUTPUT.
b: PULSE TRAINS PRODUCED BY SHAPERS.
c: PITCH DISPLAY CONTAINING ERRORS.
d: PITCH DISPLAY, ERRORS SUPPRESSED.

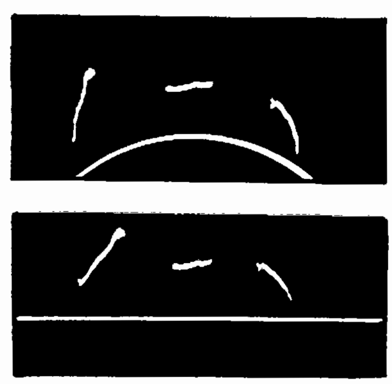
S E A SH O RE

FIG.3 SAME AS FIG.2 EXCEPT THAT PULSE GATE HAS BEEN REMOVED FROM THE CIRCUIT. NOTE SPURIOUS PULSES IN b.

(Pg14)



a
b



ONCE UPON A TIME,....

FIG.4 DISPLAY OF THE ABOVE WORDS ON A CONTINUOUS TIME BASE.
a: AMPLITUDE PATTERNS.
b: PITCH PATTERNS. (Pg16)

ONE, TWO, THREE.

FIG.5 PITCH PATTERNS FORMED BY THE ABOVE WORDS:
a: ON A CONTINUOUS TIME BASE.
b: LINEAR, ONE-SHOT TIME BASE.
(Pg18)

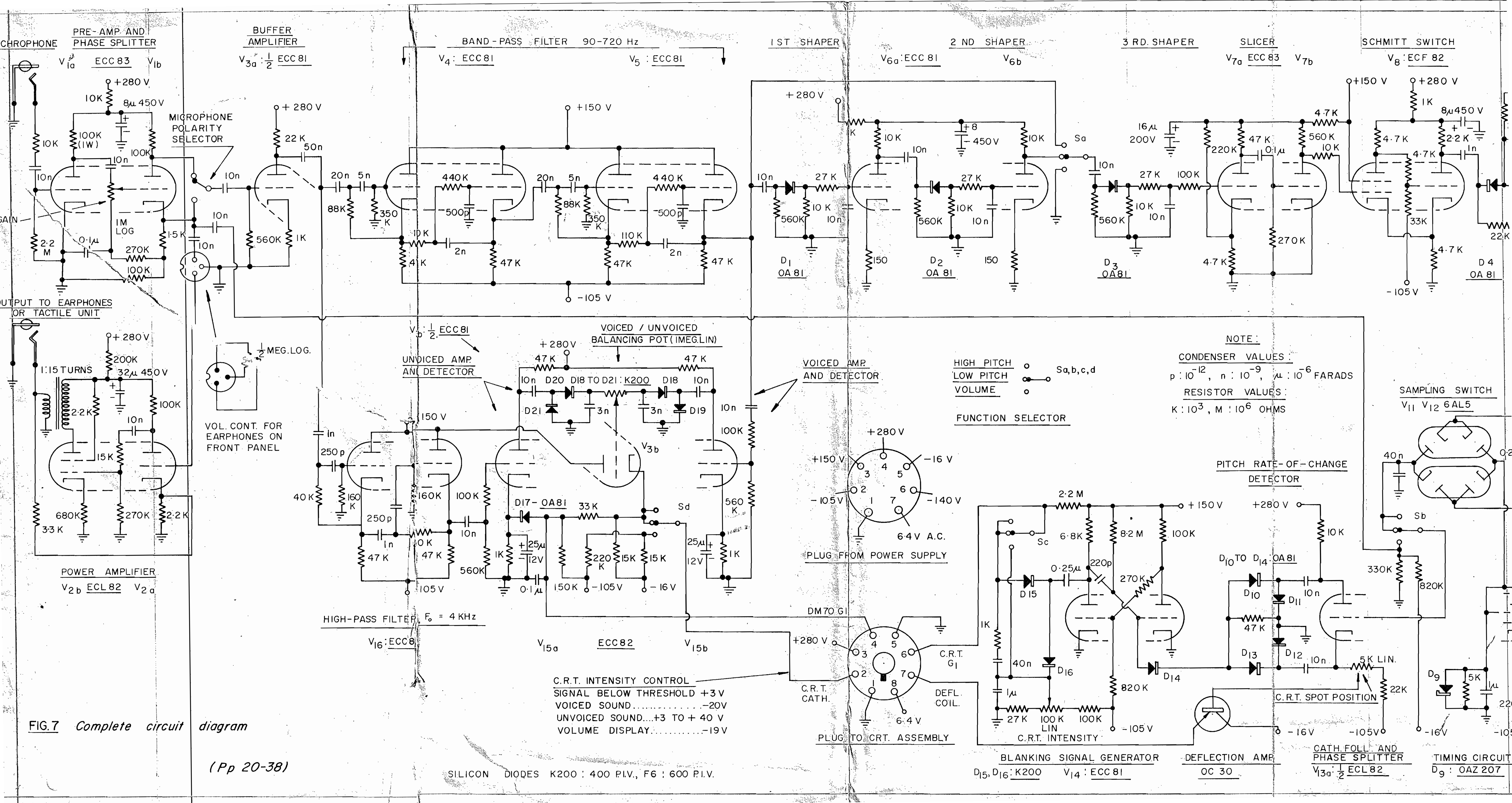


FIG. 7 *Complete circuit diagram*

(Pp 20-38)

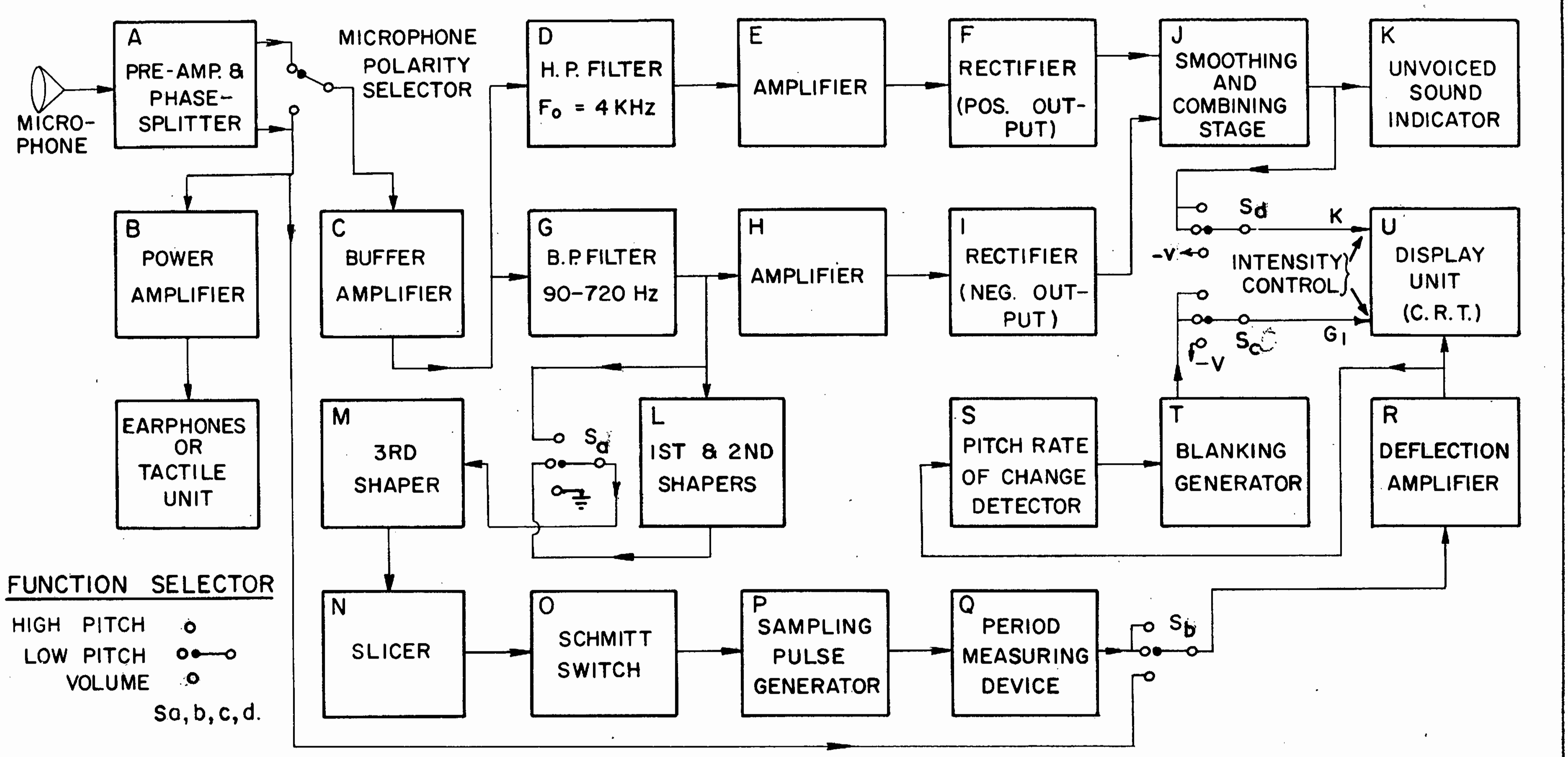


FIG. 6 *Block diagram of final design.* (Pp 19-35)

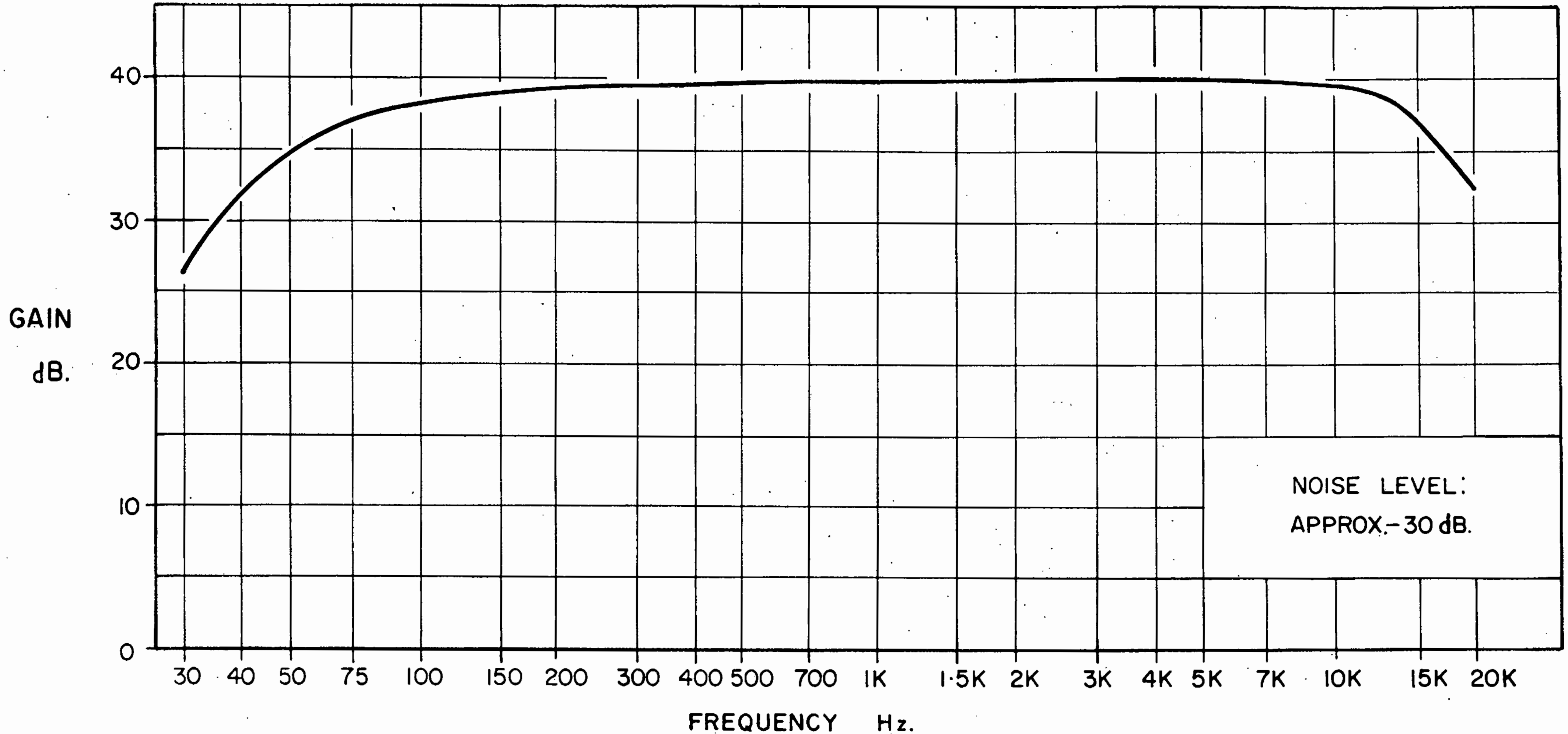
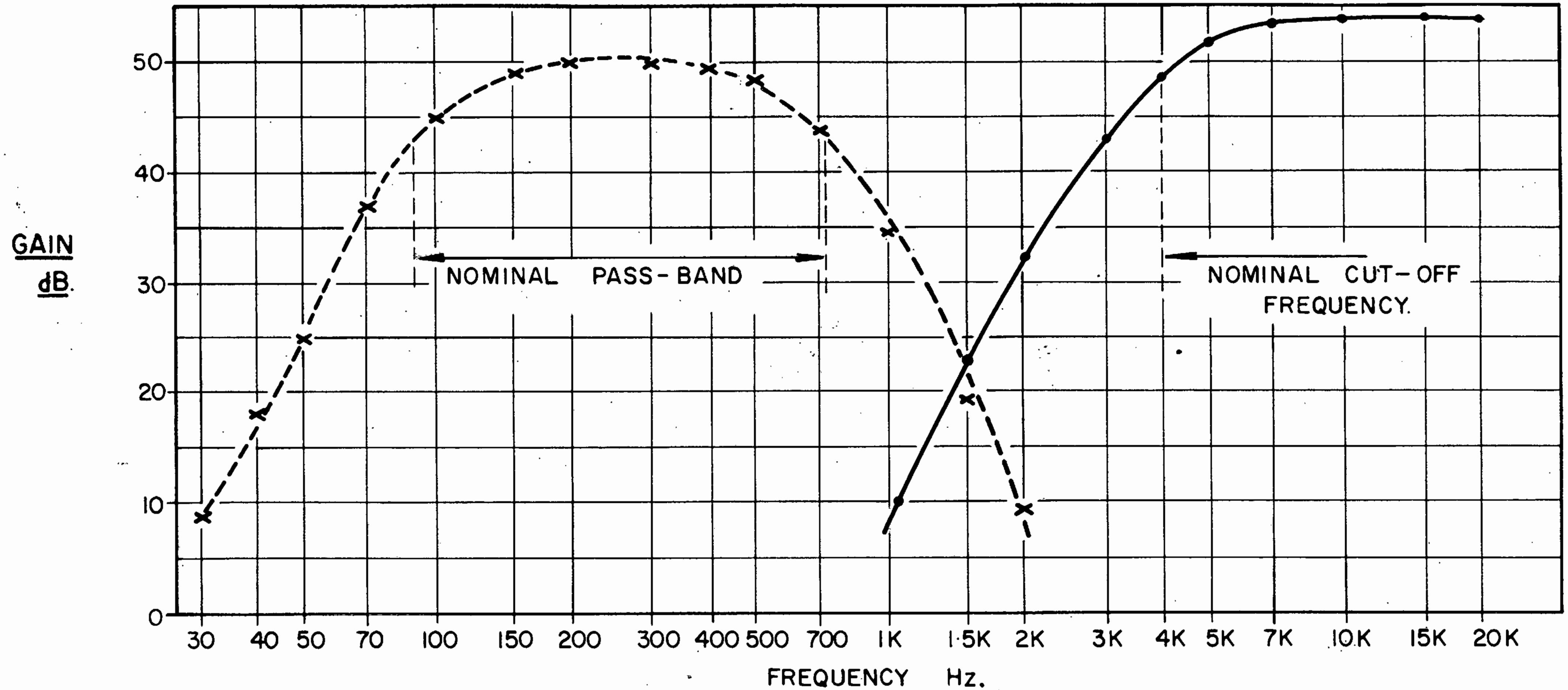


FIG. 8. *Transfer Characteristic: Microphone Input Socket to 12 Ω output Socket. Obtained at 42 mV R.M.S. Input; Both Volume Controls set at Maximum.*



TRANSFER CHARACTERISTICS: MICROPHONE SOCKET TO FILTER OUTPUT LINES.

FIG. 11. *High – Pass Filter.* ———. FIG. 12. *Band – Pass Filter.* – – – –. *(Pg. 22.)*

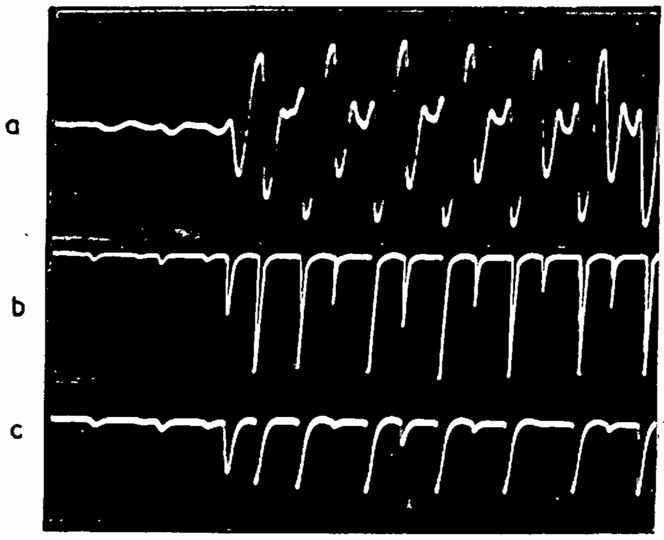
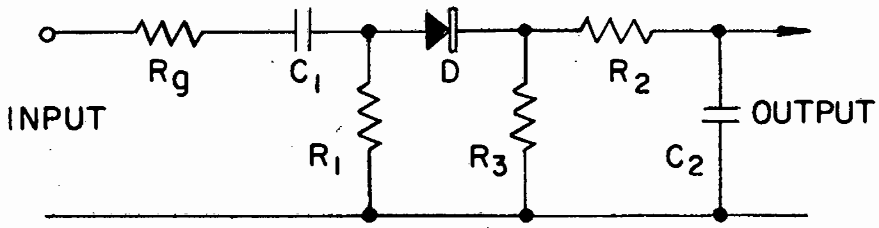


FIG. 13. WAVEFORM SHAPING. PART OF THE SOUND U AS IN POOL.
a: OUTPUT OF BAND-PASS FILTER.
b: WAVEFORM AFTER DIFFERENTIATION AND PEAK-DETECTION.
c: WAVEFORM b PARTIALLY INTEGRATED.

(Pg 24)



$R_1 = 560\ K$; $R_2 = 10\ K$; $R_3 = 27\ K$; $Rg < 10\ K$.

$C_1 = 0{\cdot}01\mu$; $C_2 = 0{\cdot}01\mu$; $D = OA\ 81$.

FIG. 14. *One Shaper Stage.* *(Pg. 24)*

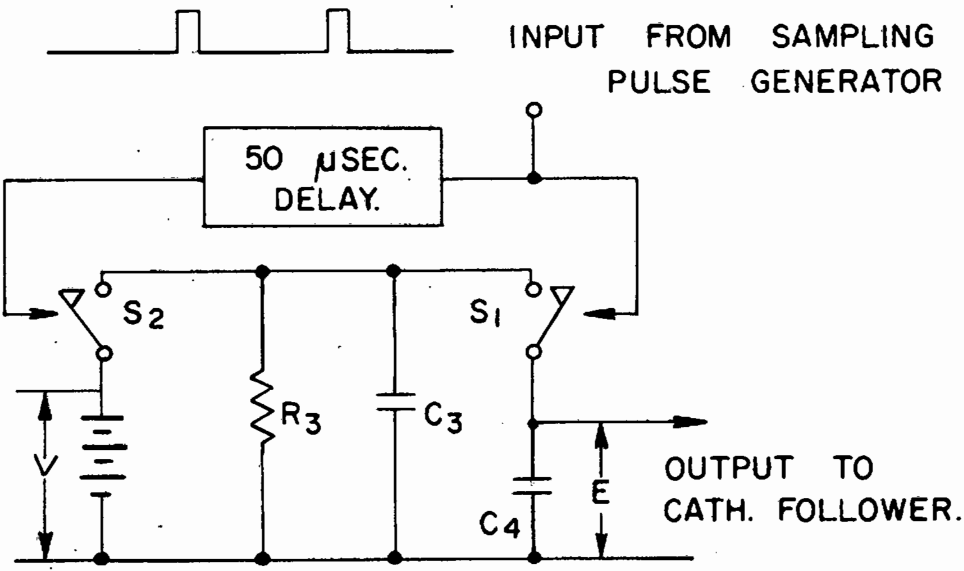


FIG. 15. *Simplified Diagram of Period-Measuring Device.* *(Pg. 26)*

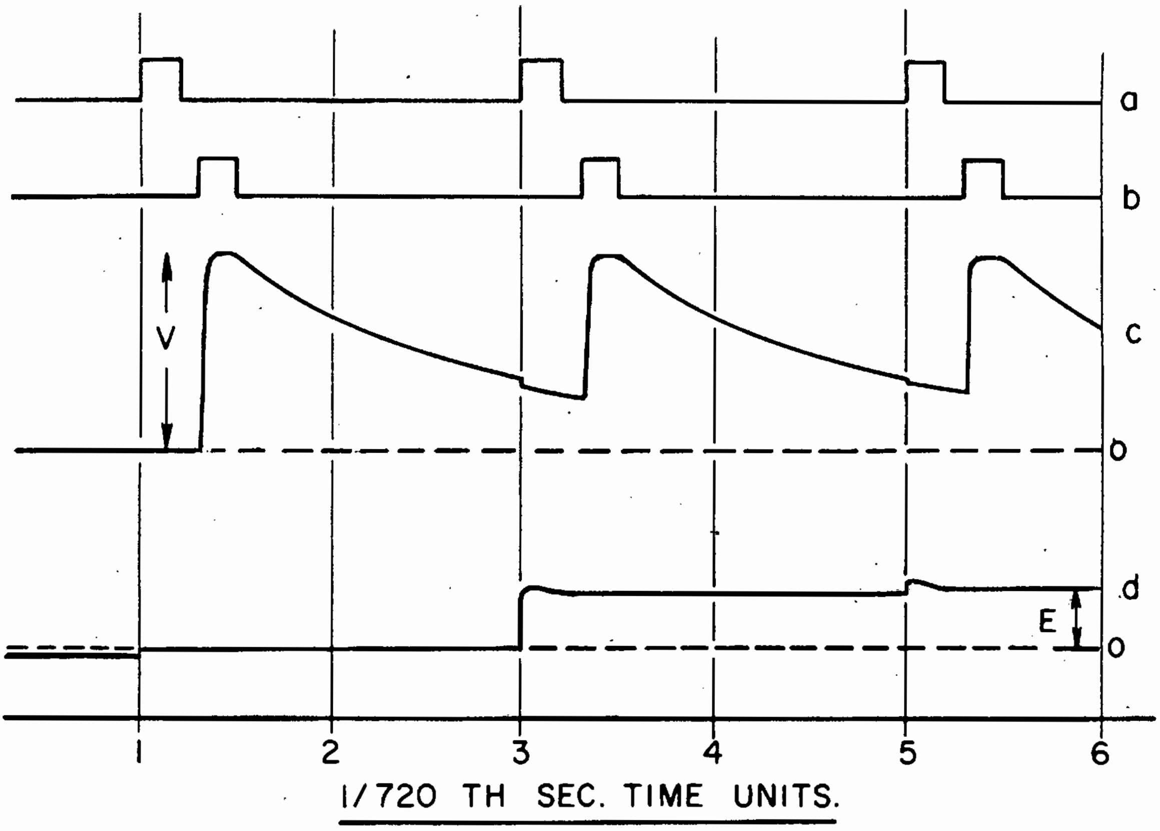


FIG. 16. *Operation of Period Measuring Device.*
a: Sampling Pulses b: Charging Pulses.
c: Waveform across Timing Network R_3C_3
d: Output Signal across C_4 (See Fig. 15.)

Pitch = 360 Hz *(Pg 26)*

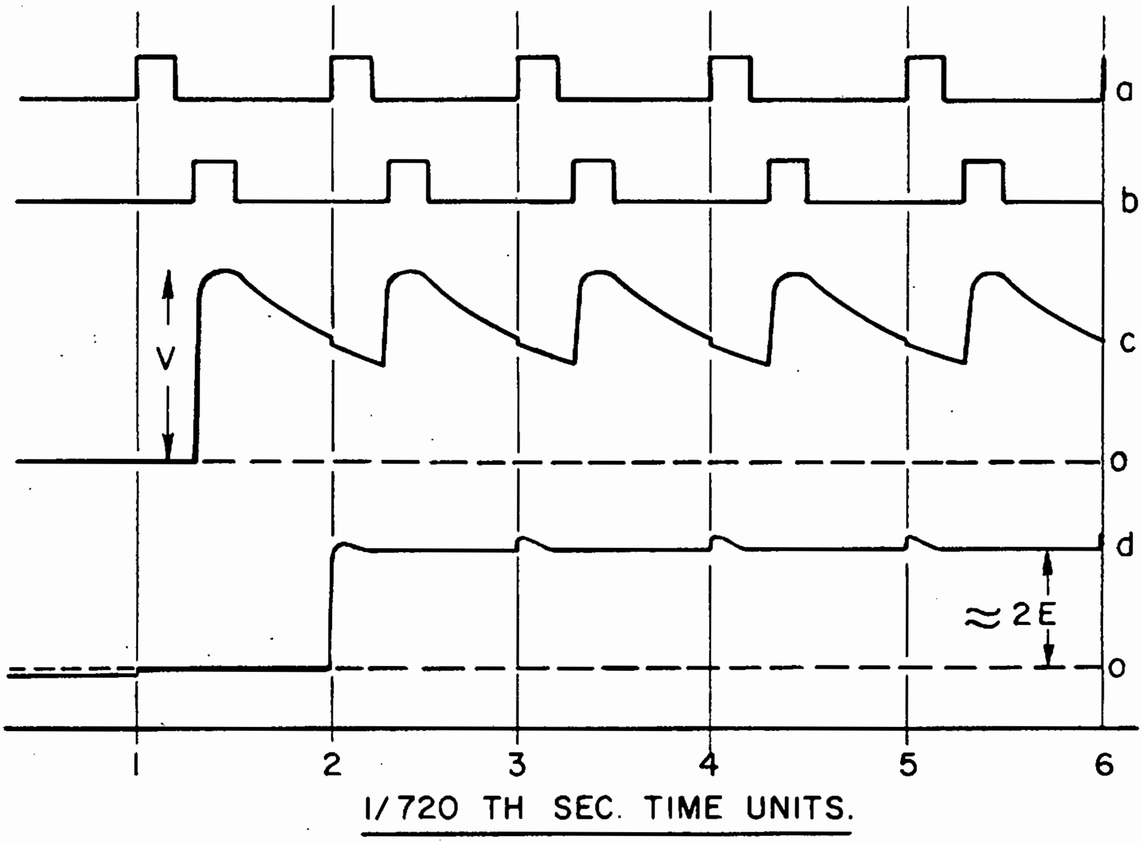


FIG. 17. *Operation of Period Measuring Device.*
For Particulars see Fig. 16.

Pitch = 720 Hz *(Pg. 26.)*

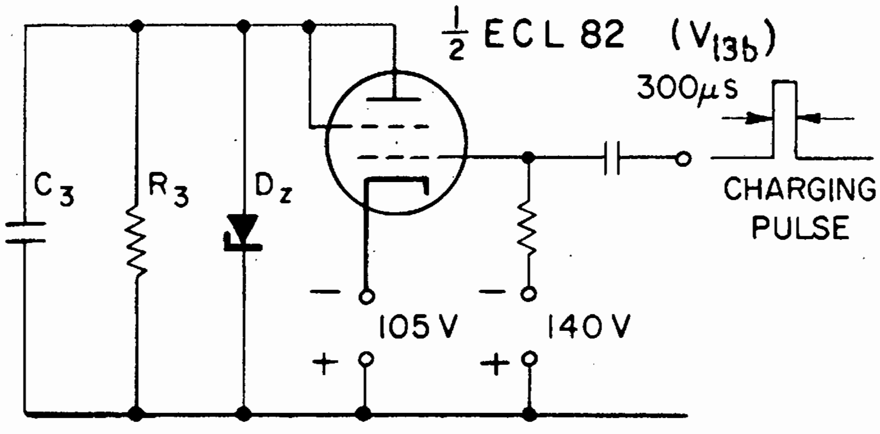


FIG. 18. *Timing Circuit* (S_2) (Pg 28)

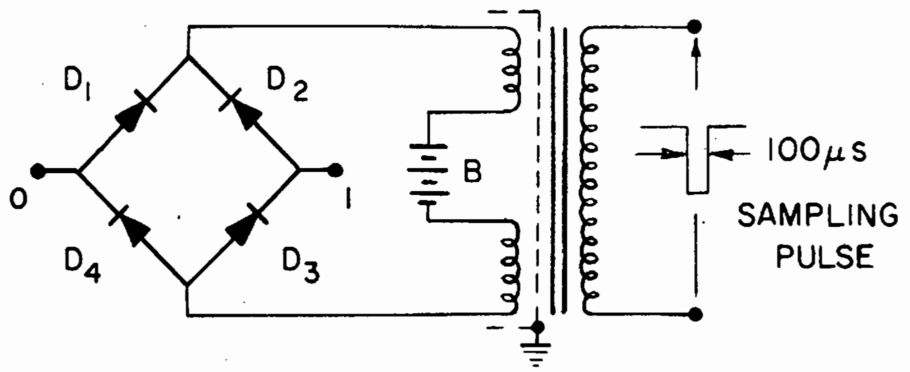


FIG. 19. *Four - Diode Switch* (S_1) (Pg 29)

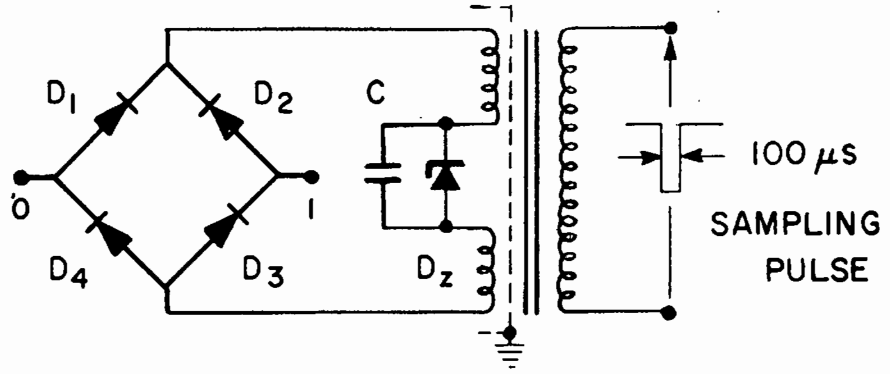


FIG. 20. *Practical 4 - Diode Switch* (S_1) (Pp 29,30)

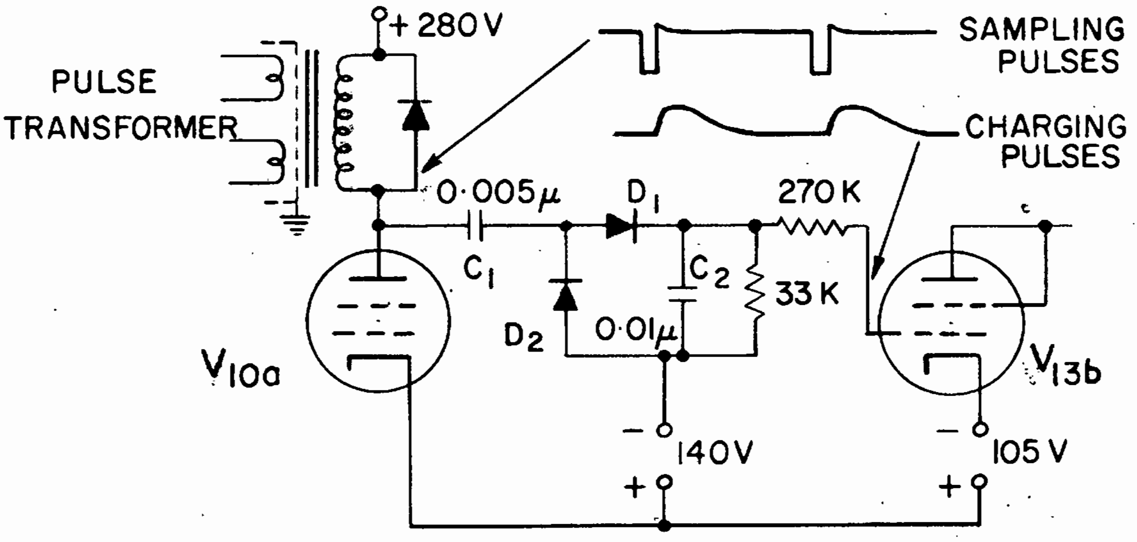


FIG. 21. *Pulse Delay Circuit* (Pg 30)

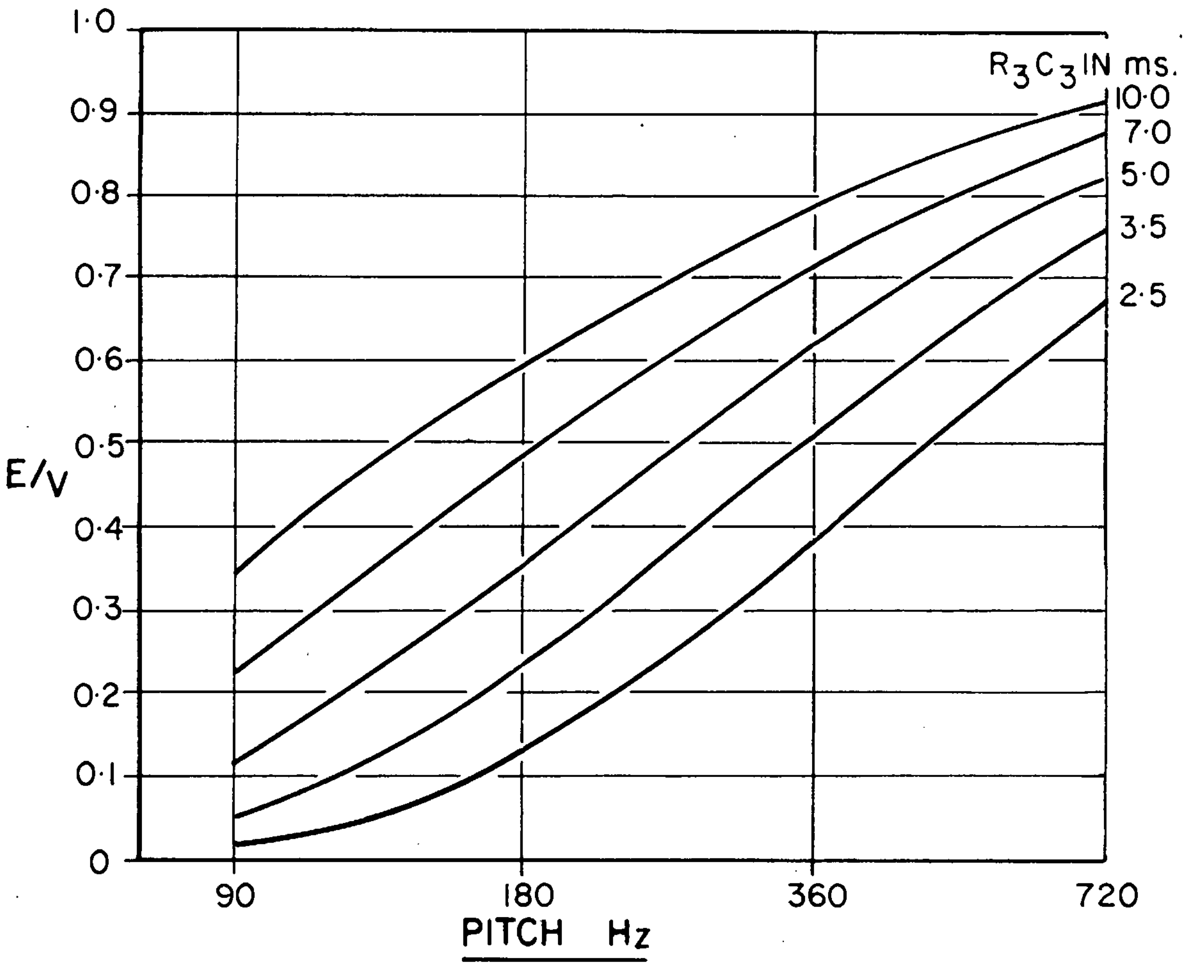


FIG. 22 *Output of Period Measuring Device vs. Pitch for 5 Values of R_3C_3 in Fig. 15. (Pg31)*

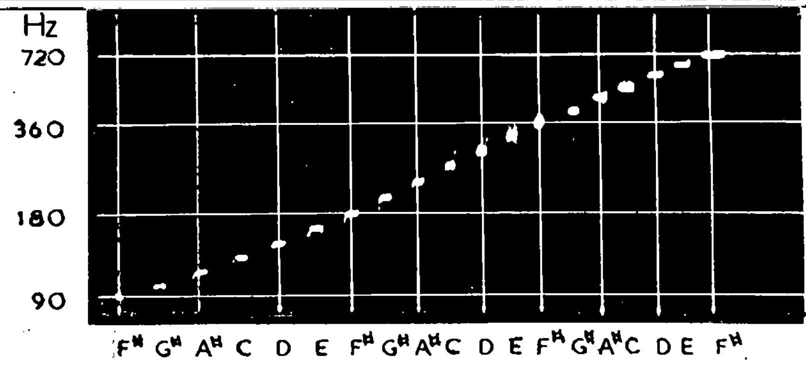


FIG. 23 PITCH DISPLAY OF A WHOLE-TONE SCALE COVERING THE THREE OCTAVES FROM 90 TO 720 Hz PLAYED UPON A MUSICAL INSTRUMENT. NOTE THAT THE FREQUENCY SCALE IS APPROXIMATELY LOGARITHMIC.

(Pg 31)

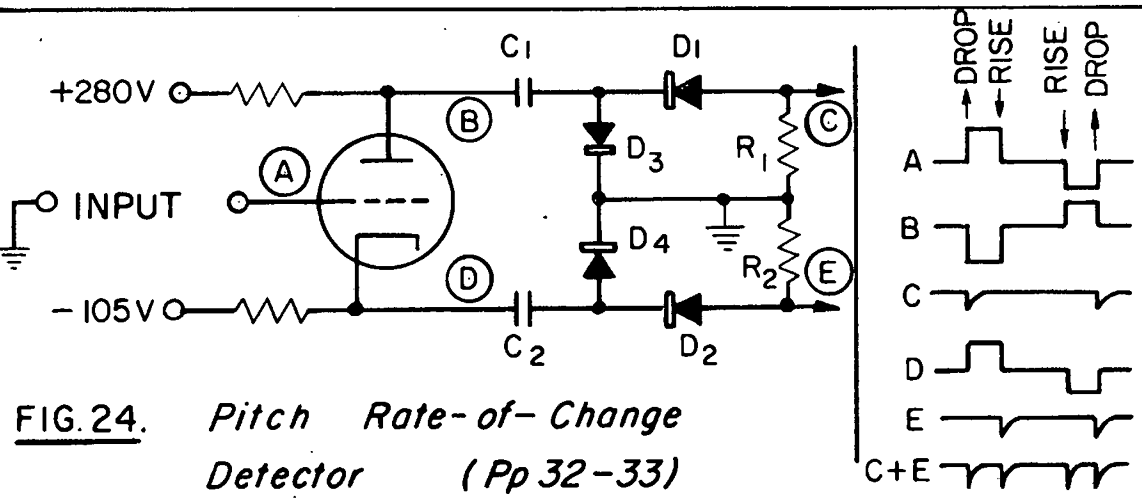


FIG. 24. *Pitch Rate-of-Change Detector (Pp 32-33)*

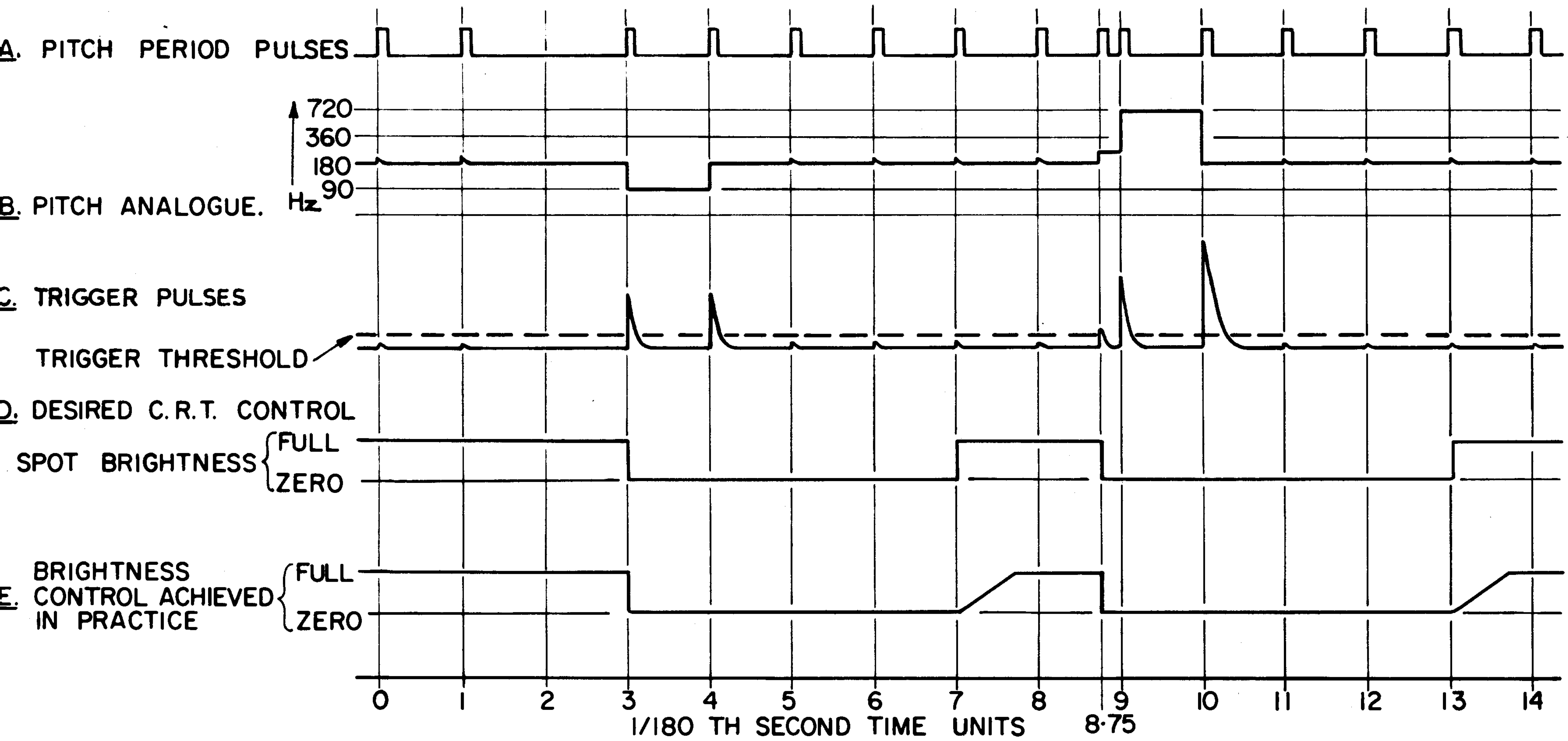


FIG. 25. *Operation of Error Detector.* (Pg. 33)

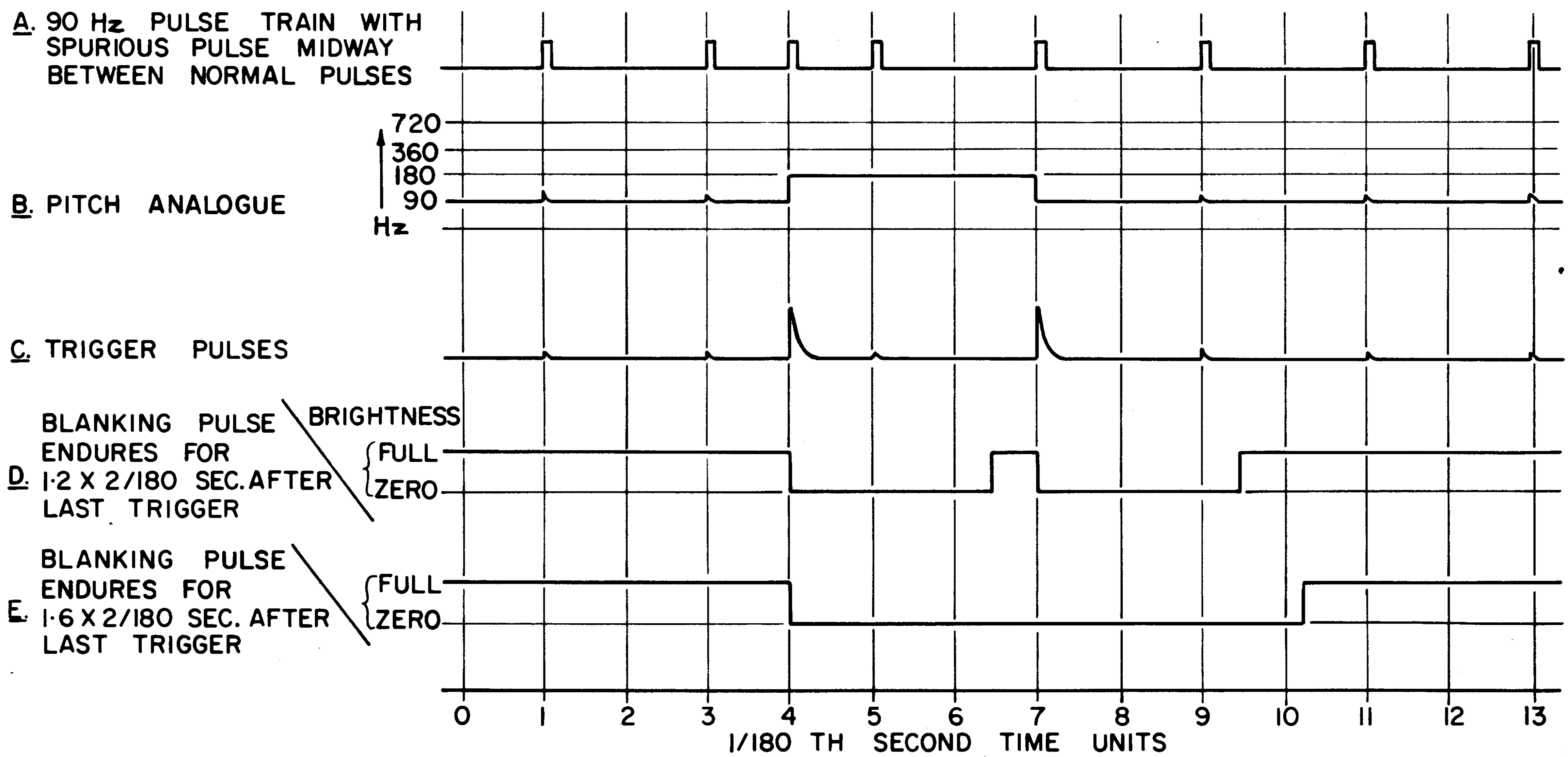


FIG. 26. *Determination of Blanking Pulse Duration.* (Pp.33,34)

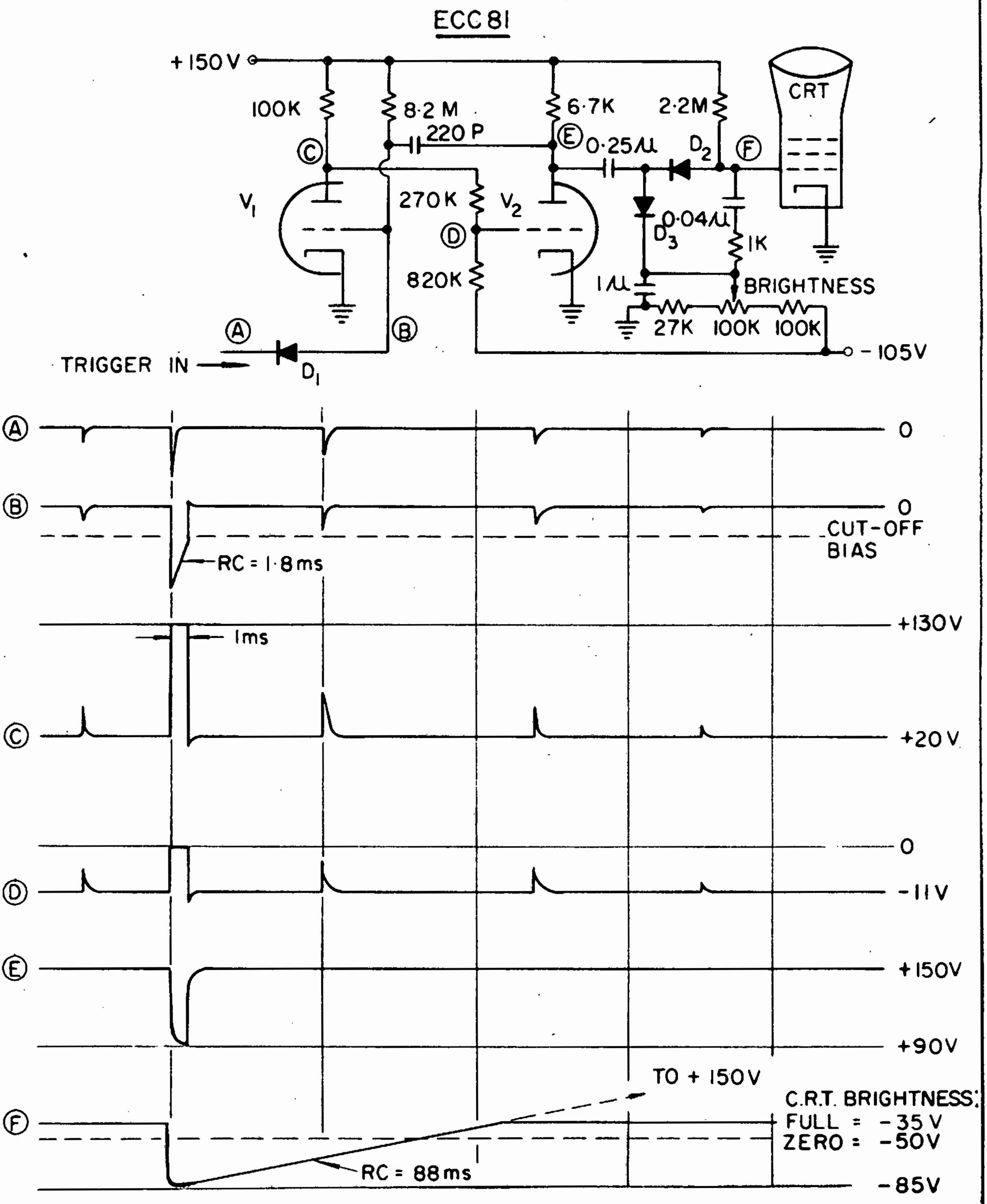


FIG. 27. *Blanking signal generator* (Pp 34,35)

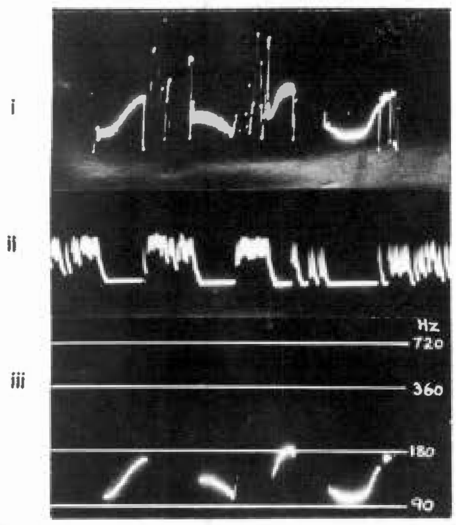


FIG. 28a. WHAT IS HIS NAME?
MALE VOICE

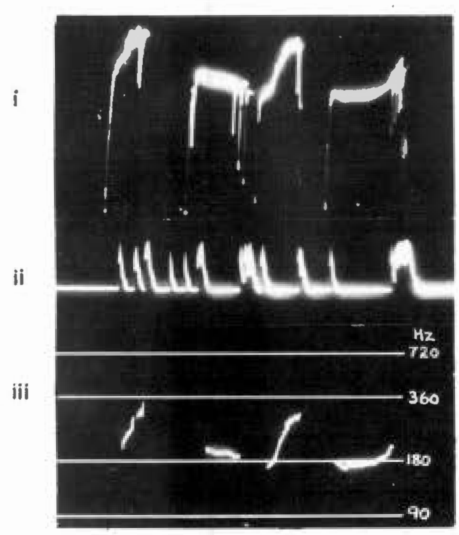


FIG. 28b. WHAT IS HIS NAME?
FEMALE VOICE

i: PITCH DISPLAY WITH ERRORS.
ii: ERROR BLANKING SIGNAL.
iii: CORRECTED PITCH DISPLAY ON A SUPERIMPOSED FREQUENCY SCALE.

COMPARE PITCH PATTERNS PRODUCED BY MALE AND FEMALE VOICES, NOTING SIMILARITY OF FORM.

(Pg 35)

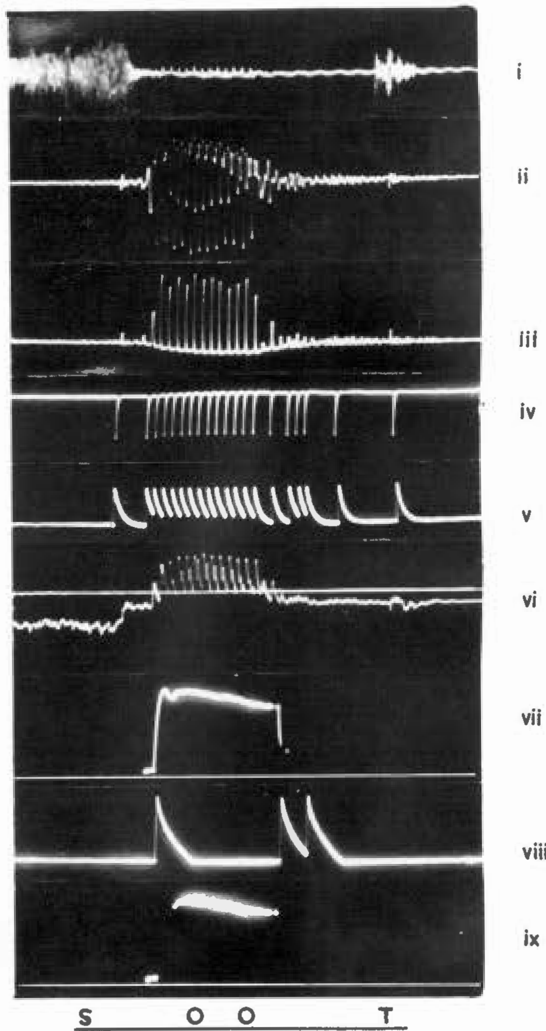


FIG. 28c. WAVESHAPES PRODUCED BY THE WORD 'SOOT' AT VARIOUS POINTS IN THE CIRCUIT.

i: OUTPUT OF HIGH-PASS FILTER.
ii: OUTPUT OF LOW-PASS FILTER.
iii: OUTPUT OF SHAPER.
iv: CHARGING PULSES IN PERIOD MEASURING DEVICE.
v: SIGNAL ACROSS TIMING NETWORK (C_3R_3).
vi: C.R.T. BRIGHTNESS CONTROL.
vii: PITCH DISPLAY WITH ERRORS.
viii: ERROR BLANKING SIGNAL AT C.R.T. GRID.
ix: CORRECTED PITCH DISPLAY.

NOTE: NEGATIVE POLARITY IS UPWARDS.

(Pg 38)

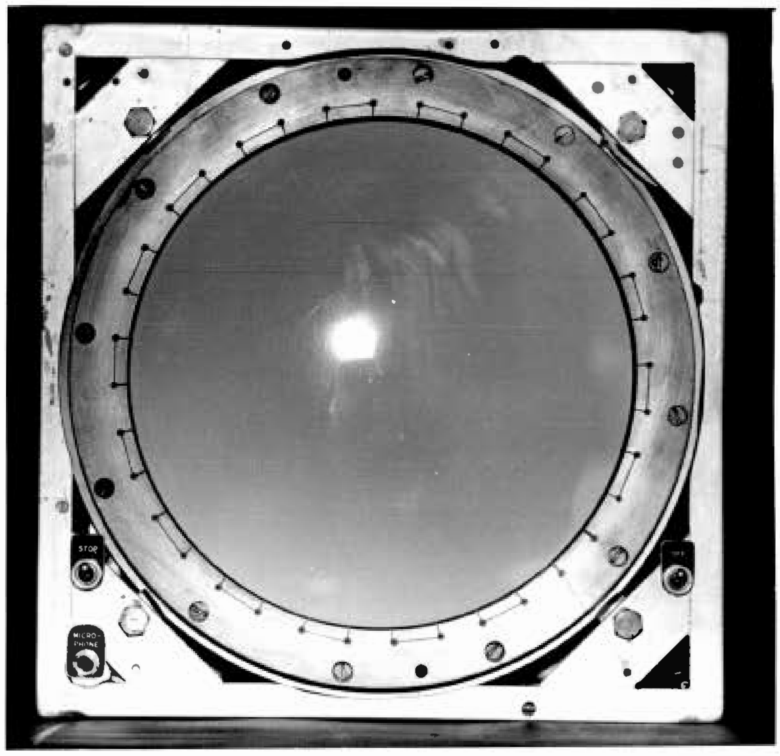


FIG 29. *Front view showing cathode ray tube inside rotating ring supported in framework*

(Pg 39)

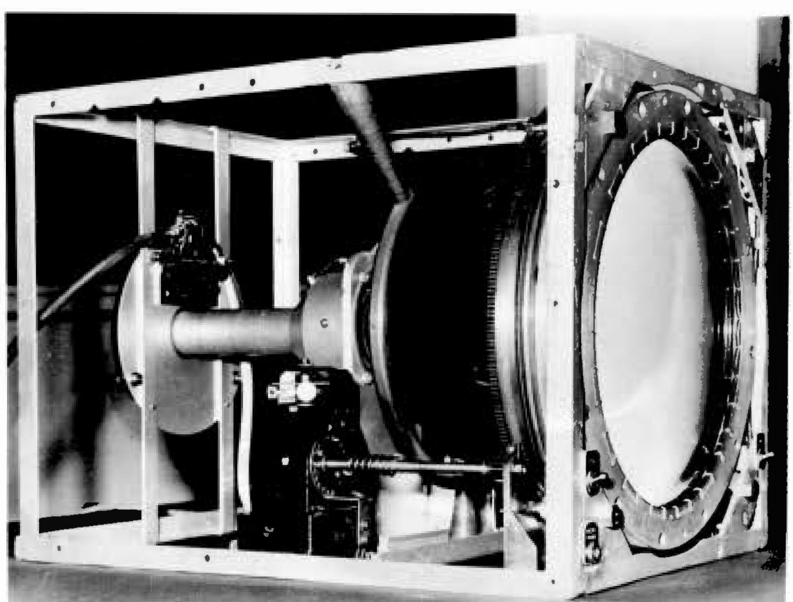


FIG. 30. *Oblique side view showing face of C.R.T. supported in ring-gear and neck of tube entering slip-ring assembly.*

(Pg 39)

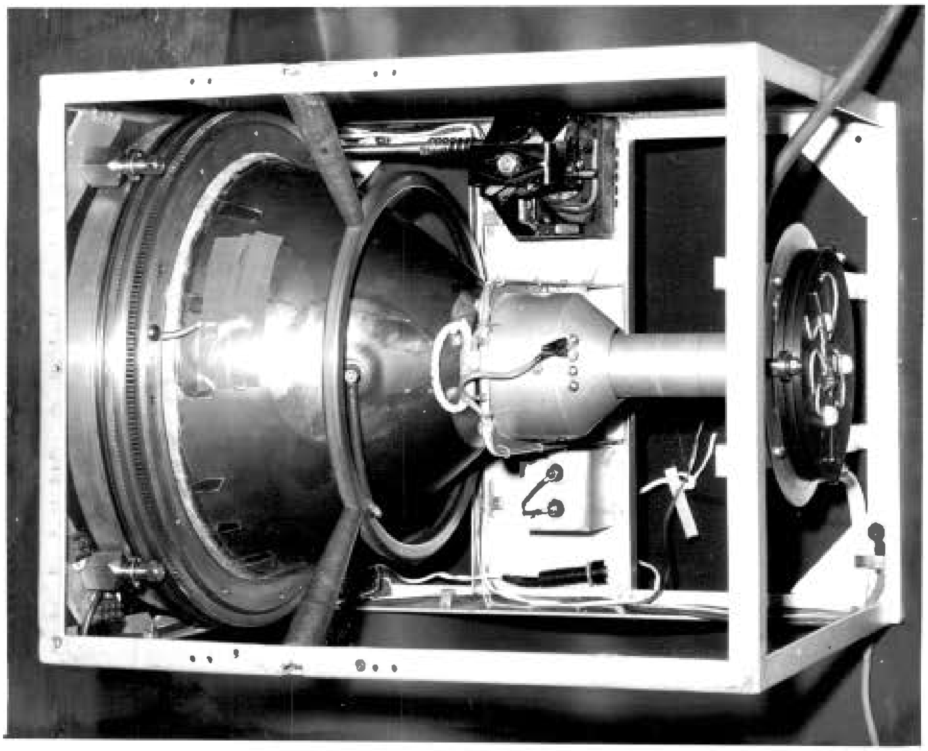


FIG 31 *Top view showing two of the four ball bearings mounted on pillars carrying the C.R.T. Note earth strap between outer conductive coating and ring-gear, and E.H.T. ring suspended on insulating pillars. The E.H.T. collector saddle is visible under this ring. Deflecting coil is inside magnetic screening cup over neck of tube.*

(Pg 39)

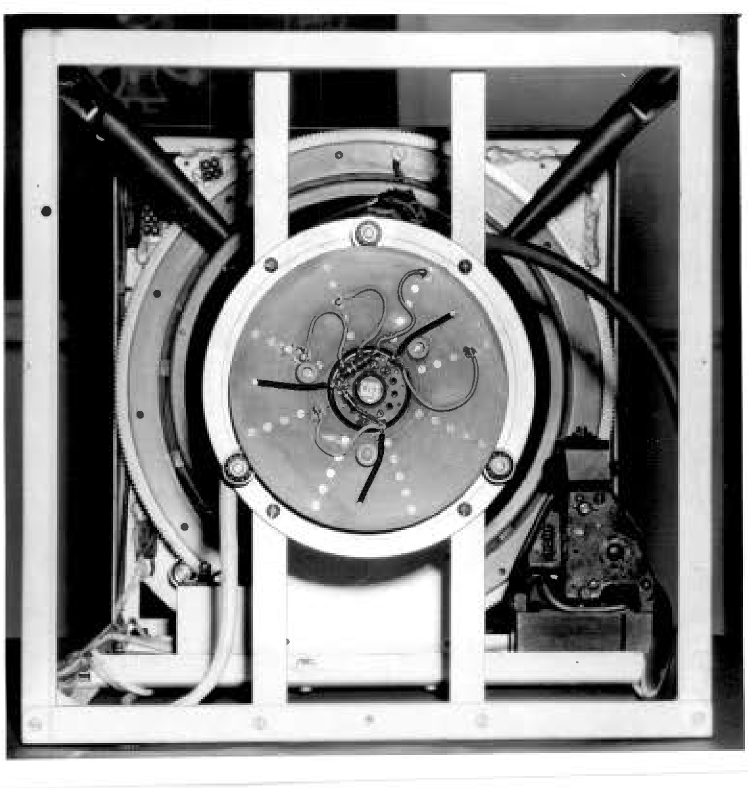


FIG 32 *Back view showing insulated disc supporting tube neck and carrying 5 slip rings. Note ball bearings around periphery of disc.*

(Pg 39)

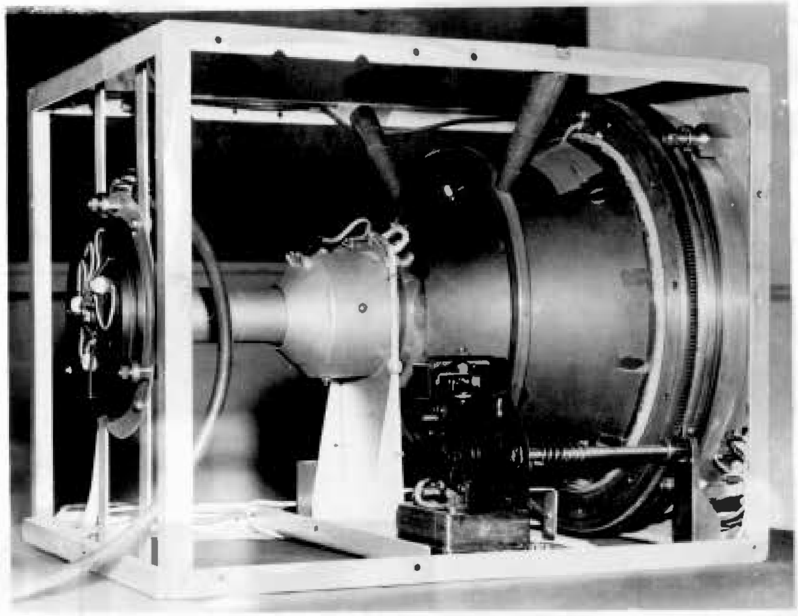


FIG 33 *Side view showing earthing strap at top of ring gear, woolen jacket in position around tube, driving motor in the centre foreground and flexible driving shaft between motor and pinion.*

(Pg 39)

FIG 34 *Top chassis: power supply (left) and E.H.T. generator (right)*
Lower chassis: complete electronic assembly.

(Pg 40)

FIG 35 *Oblique back view of assembled instument showing the two chassis in position above and below tube neck. The tactile unit is seen at the left* (Pg 40)

FIGURES 36, 37, 38 : *See next two pages*

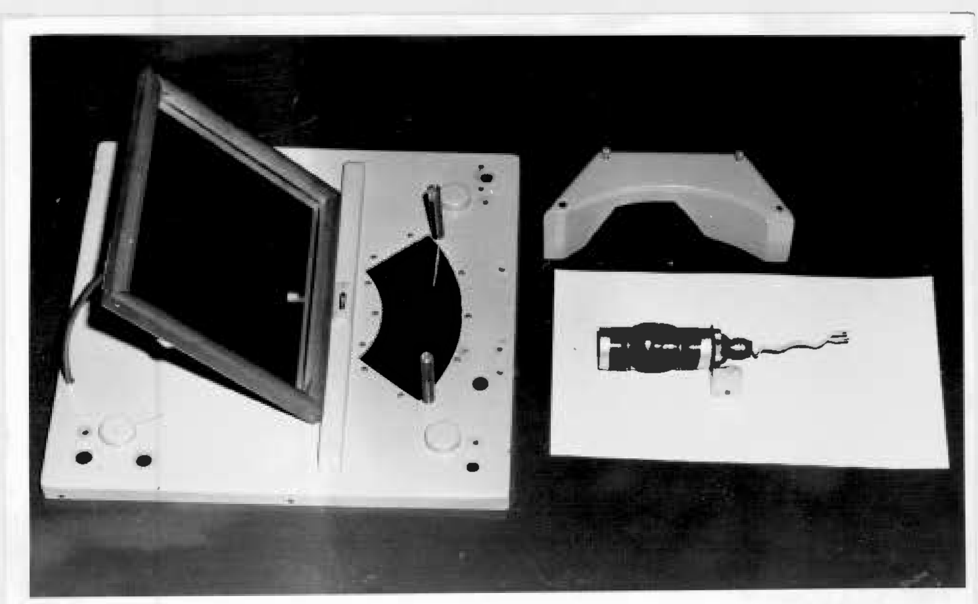


FIG 39 *Left: Front cover with mirror and viewing aperture. Unvoiced-sound indicator is in slot next to mirror. Right: Lamphouse (top) and lamp inside light shield with aperture which is covered by infra-red filter.* (Pg 40)

FIG 40 : *See after fig. 38.*

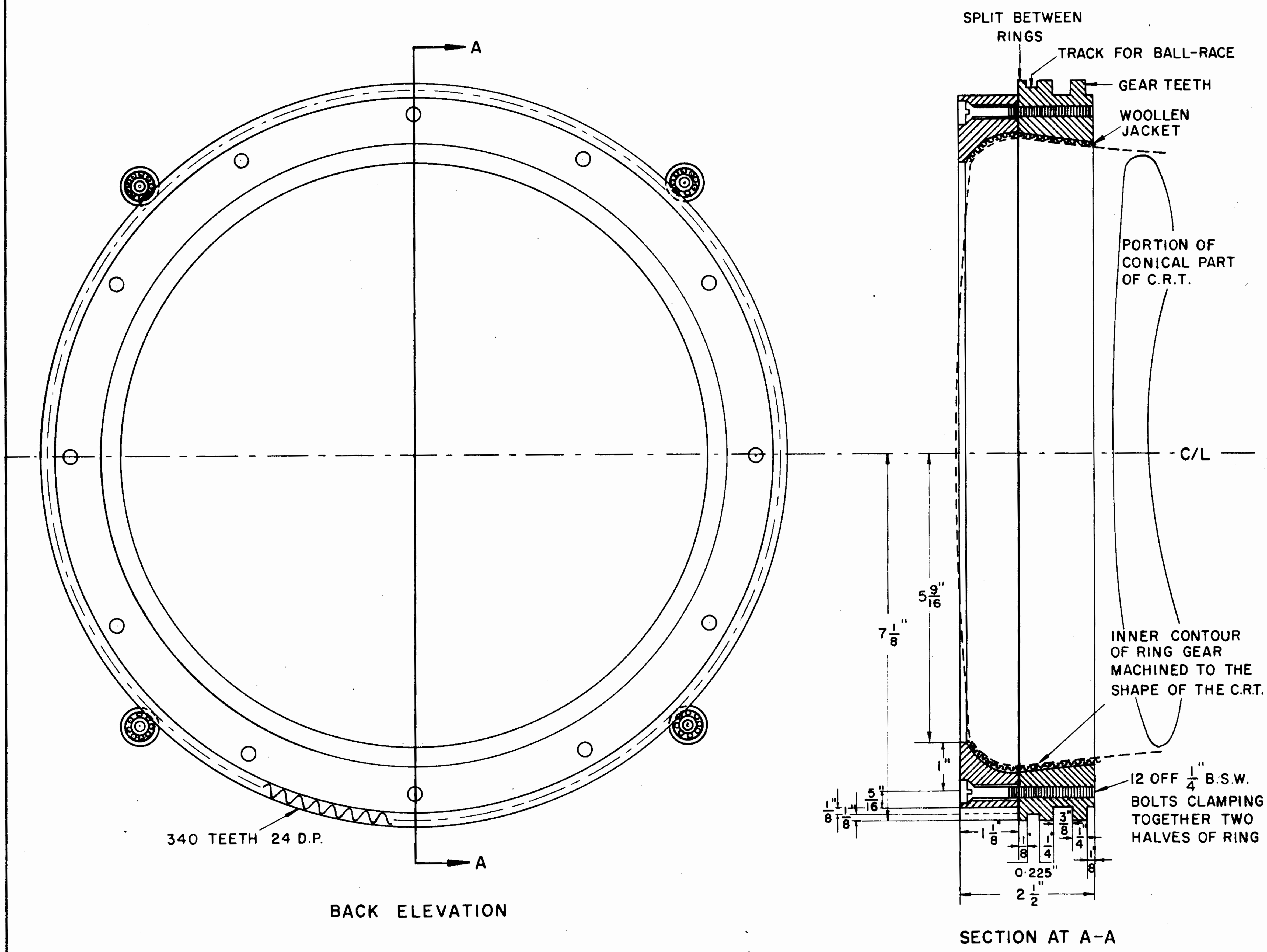


FIG. 36. *Ring gear assembly.* SCALE : $\frac{1}{2}$" = 1"

(Pg 39)

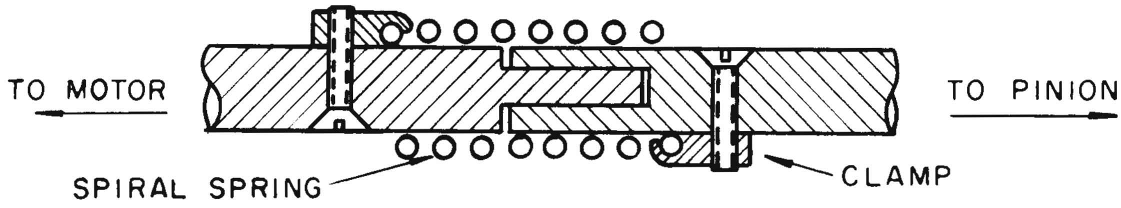


FIG. 37. *Flexible coupling.* (Pg 39)

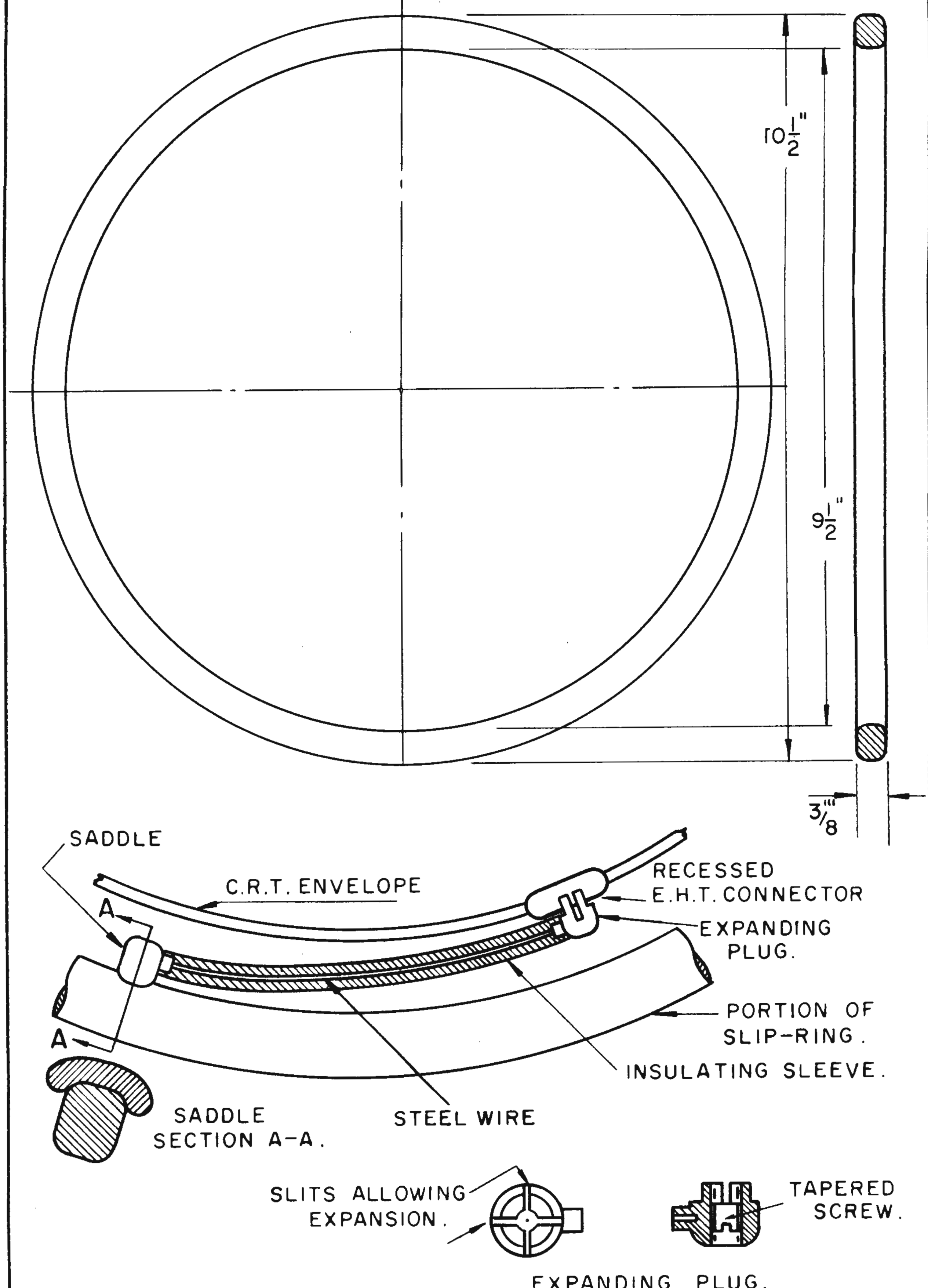


FIG. 38 *E.H.T. Slip-ring, saddle, insulated steel coupling wire and expanding plug.*

(Pg 40)

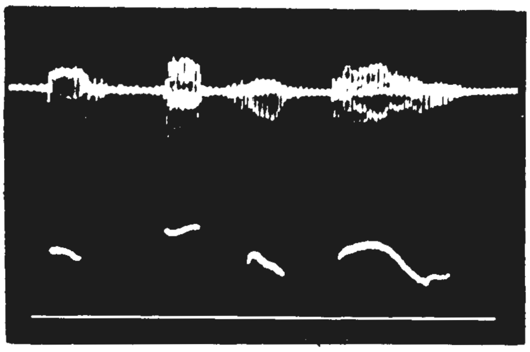
Fig. 39: Turn back 2 pages.

FIG. 40 *Front view of complete equipment showing child speaking into microphone whilst viewing patterns on the screen and feeling sound vibrations on on tactile unit*

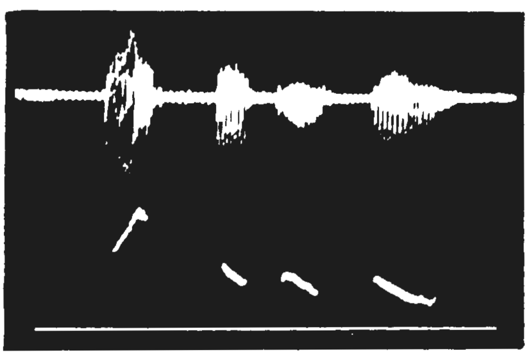
(Pg 41)

FIG 41 *A group of 3 children with their teacher. Each child is able to follow the vocal exercises of the speaker by combining the visual and tactile stimuli. Note tactile units on childrens' laps*

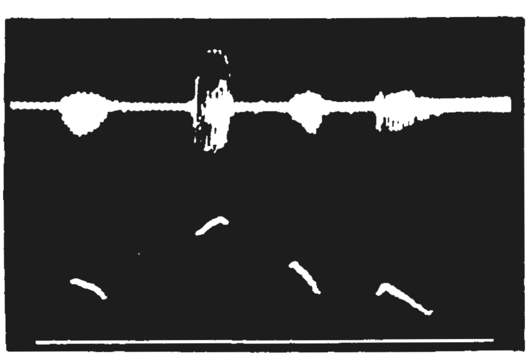
(Pg 42)



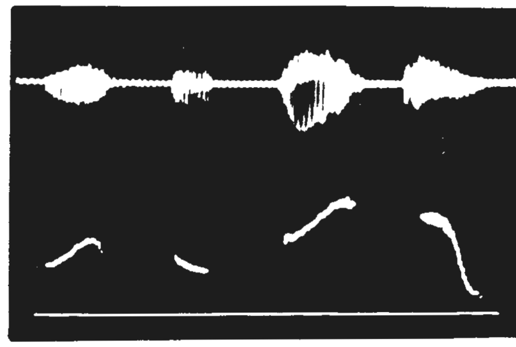
a HIS CAT IS BROWN.



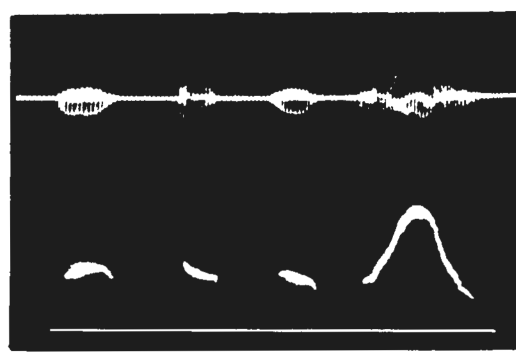
b HIS CAT IS BROWN.



c HIS CAT IS BROWN.



d HIS CAT IS BROWN.



e HIS CAT IS BROWN.

FIG. 42. FIVE VARIATIONS OF A STATEMENT DISPLAYED AS AMPLITUDE AND PITCH PATTERNS. NOTE HOW CHANGES OF MEANING, AS DIFFERENT WORDS ARE STRESSED, ARE CLEARLY ILLUSTRATED AS WELL AS THE DURATION OF WORDS AND THEIR POSITIONS ALONG THE TIME SCALE.

(Pg 46)

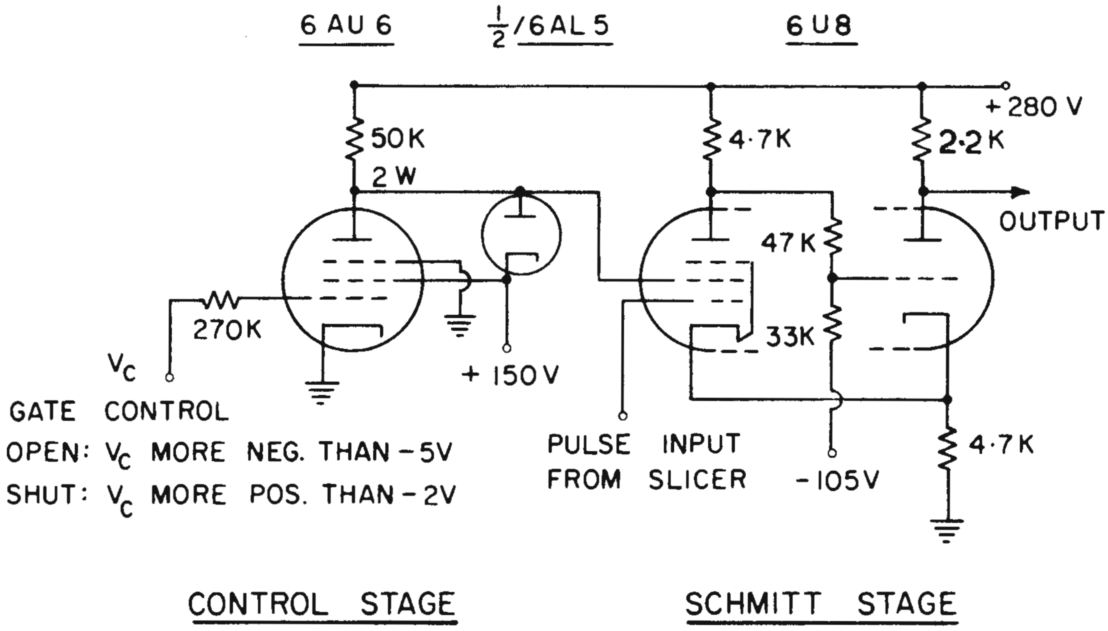


FIG. A1. *Pulse gate. Condition of gate is determined by voltage on g_2 of the 6U8.* (Pg 14)

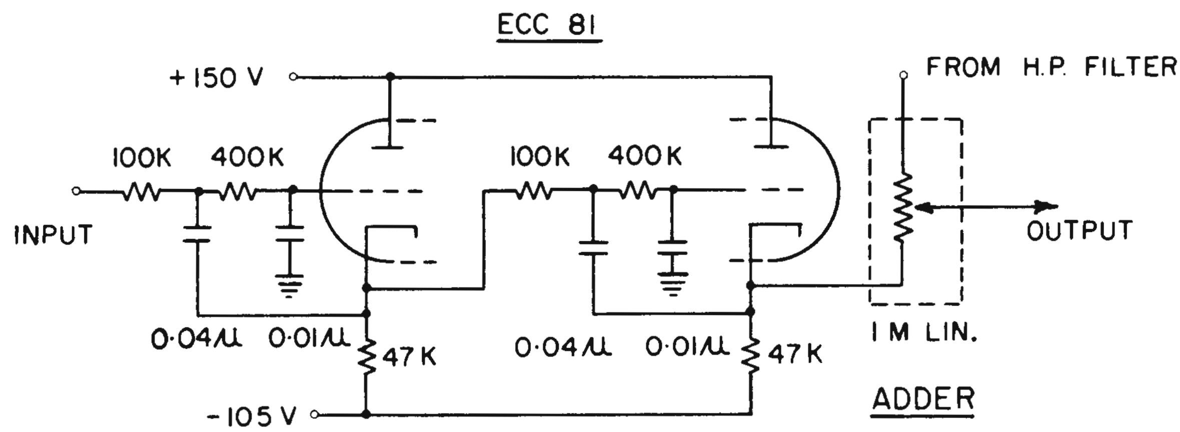


FIG. A2. *Low pass filter. F_0 = 50 Hz. Slope beyond F_0: 24 dB/octave.* (Pg 15)

FIG. A3. *See next page*

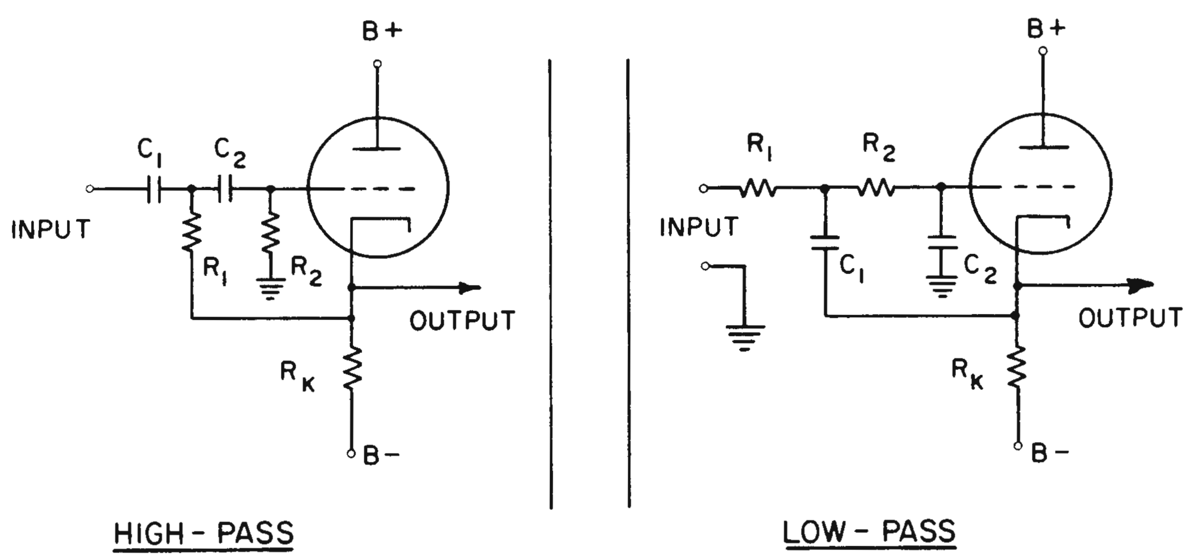


FIG. A4. *R-C active filters.* (Pp 21,22, Ai)

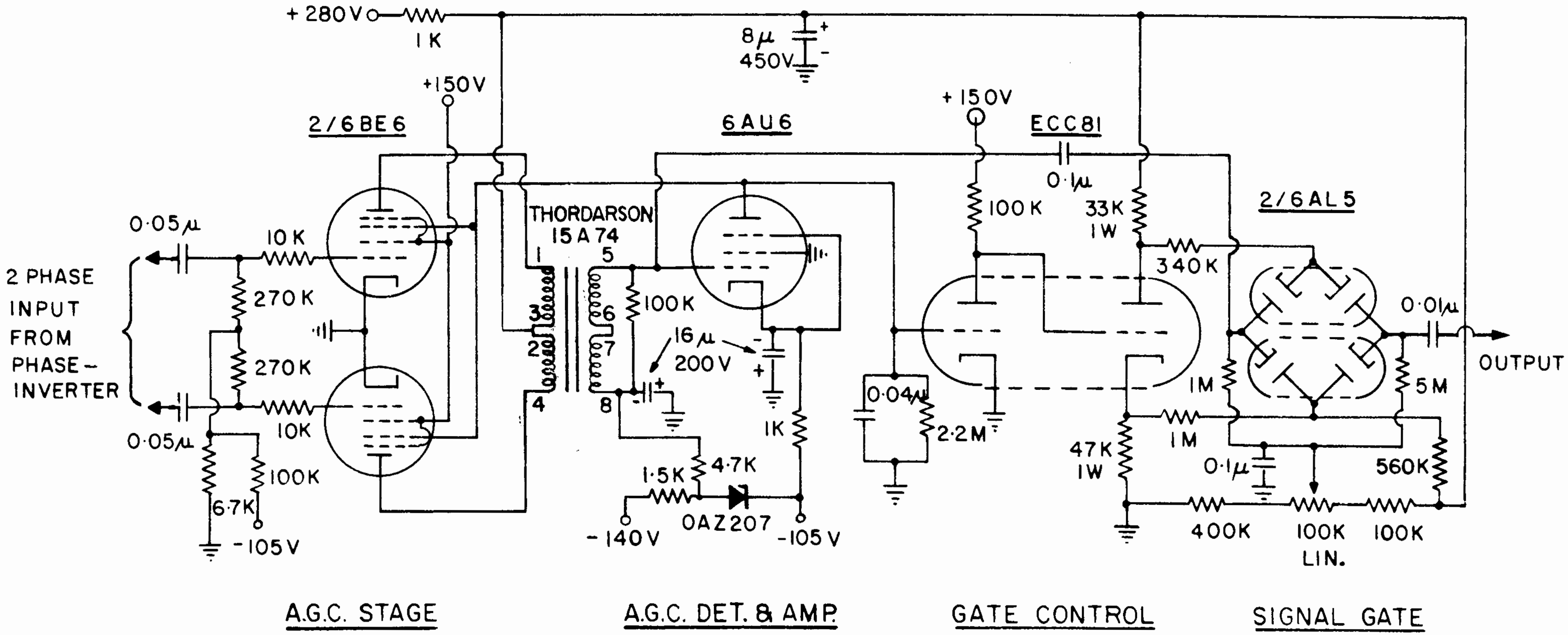


FIG. A3. *Gain-controlled amplifier and signal gate* (Pp 15,16)

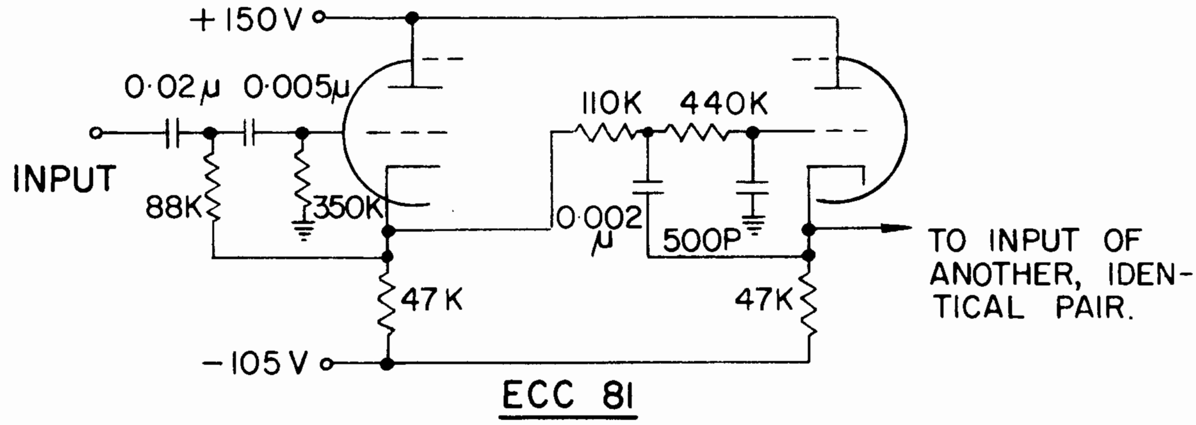


FIG. A5. *Bandpass Filter. 90 to 720 Hz (Pp 22, Ai)*

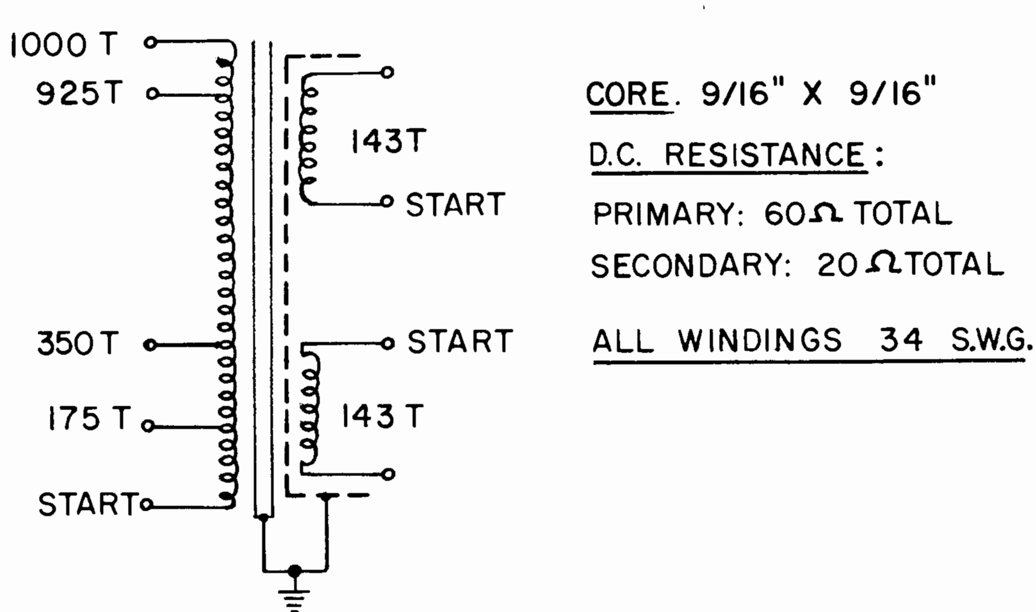


FIG. A6. *Pulse Transformer (Pp Aiii, Aiv)*

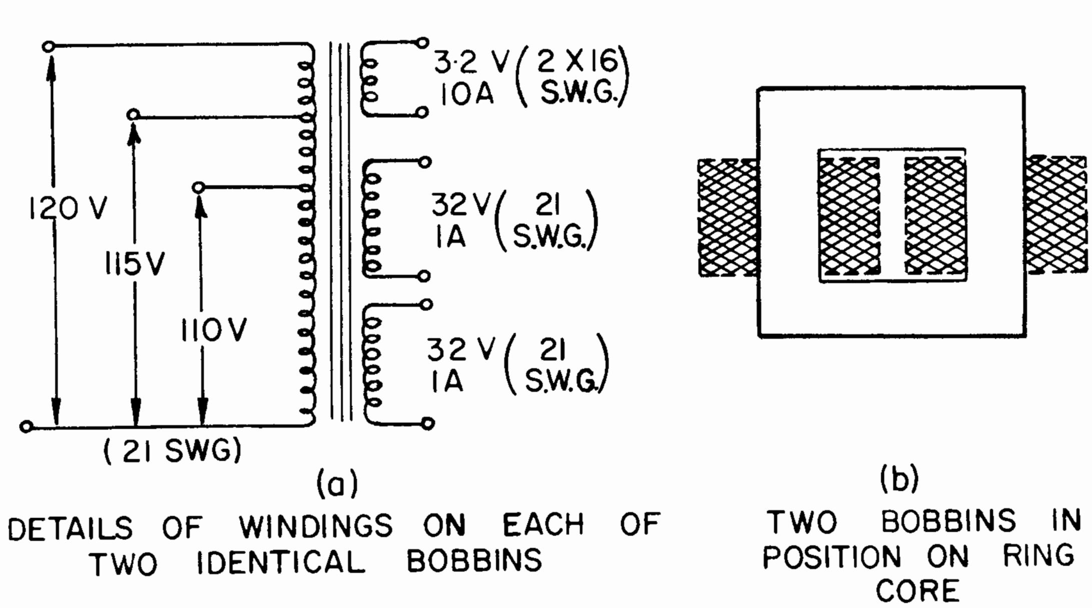


FIG. A7. *Power Transformer (Pg 38)*

REFERENCES.

1. Anderson F., J. Acoust. Soc. Am., Vol. 32, No. 8 (August 1960) pp 1065 - 1074 : "An experimental Pitch Indicator for Training Deaf Scholars."

2. Dolanski L. O., J. Acoust. Soc. Am., Vol. 27, No. 1 (January 1955) pp 67 - 72 : "An Instantaneous Pitch-Period Indicator."

3. Kallenbach W., Akustica, Akust. Beih. 1 (1951) pp 37 - 42 : "Eine Weiterentwicklung des Tonhöhenschreibers mit Anwendungen bei phonetischen Untersuchungen."

4. Kallenbach W., Frequenz, Band 16, No. 2 (Februar 1962) pp 37 - 42 : "Die Untersuchung der Sprache mit dem Tonhöhenschreiber."

5. Grützmacher M. und Lottermoser W. Akust. Z. 2 (1937) pp 242 - : "Über ein Verfahren zur trägheitsfreien Aufzeichnung von Melodiekurven."

6. Reisz R. R. and Schott L., J. Acoust. Soc. Am., Vol.18, No. 1 (July 1946) pp 50 - 61 : "Visible Speech Cathode Ray Translator."

7. Plant G. R. G., The Teacher of the Deaf, Vol. 58, No. 343 (February 1960).

8. Gold B., J. Acoust. Soc. Am., Vol. 34, No. 7 (July 1962) pp 916 - 921 : "Computer Program for Pitch Extraction."

9. Harris C. M. and Weiss M. R., J. Acoust. Soc. Am., Vol.35, No. 3 (March 1963) pp 339 - 343 : "Pitch Extraction by Computer Processing of High-Resolution Fourier Analysis Data."

10. Noll A. M., J. Acoust. Soc. Am., Vol. 36, No. 2 (February 1964) pp 296 - 302 : "Short-time Spectrum and 'Cepstrum' Techniques for Vocal Pitch Detection."

11. Weiss M. R. and Harris C. M., J. Acoust. Soc. Am., Vol. 35, No. 2 (February 1963) pp 207 - 214 : "Computer Technique for High-Speed Extraction of Speech Parameters."

12. Gill J. S., Nature Vol. 189, No. 4759 (January 14, 1961) pp 117 - 119 : "A Versatile Method for Short-Term Spectrum Analysis in 'Real-Time'". For a synopsis of this paper see: Wireless World Vol. 67, No. 3 (March 1961) pp 112 - 113 : "A.F. Spectrum Analyser..."

13. Wood D. E. and Hewitt T. L., J. Acoust. Soc. Am., Vol. 35, No. 8 (August 1963) pp 1274 - 1278 : "Spectrographic Pictures of Speech."

14. Harris C. M. and Waite W. M., J. Acoust. Soc. Am., Vol. 35, No. 4 (April 1963) pp 447 - 450 : "Gaussian-Filter Spectrum Analyser."

15. Huggins W. H., J. Acoust. Soc. Am., Vol. 22, No. 6 (November 1950) pp 765 - 767 : "System-Function Analysis of Speech Sounds."

16. Gruenz O.O. and Schott L.O., J. Acoust. Soc. Am., Vol. 21, No. 5 (September 1949) pp 487 - 495 : "Extraction and Portrayal of Pitch of Speech Sounds."

17. Fryer W. D., Electronics, Vol. 32, No. 15 (April 10,1959) pp 68 - 70 : "How to Design Low Cost Audio Filters."

18. Swaffield G. L., Wireless World, Vol. 64, No. 7,(July 1958) pp 344-348 : "The Schmitt Multivibrator."

Reprinted from THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, Vol. 32, No. 8, 1065–1074, August, 1960

An Experimental Pitch Indicator for Training Deaf Scholars

F. ANDERSON
Electro-Acoustic Laboratory, School for the Deaf, Worcester, Union of South Africa
(Received December 21, 1959)

An instrument is described which extracts from the complex speech wave, as produced by the deaf child or its teacher, information related to the subjective pitch of the sound. It then displays this information on the long-persistence screen of a revolving cathode-ray tube in such a manner that a continuous graph of pitch vs time is obtained. The deaf child is, therefore, enabled to compare visually the pitch patterns produced by its own voice with the standard as produced by the teacher and to correct defects in the pitch changes, rhythm, and phrasing of his speech.

An experimental model of this instrument is being used successfully, but some deficiencies exist which require further study before an improved design can be finalized.

INTRODUCTION

TEACHERS who are employed to educate children afflicted with severe deafness at an early age, before normal speech habits could have been formed, find that modern conventional methods of speech training are quite effective when dealing with the average intelligent child. Articulation of words may be taught by recourse to methods such as feeling the vibrations on the teacher's chest, throat, etc., in conjunction with observing the motions of the different speech organs. It is also found, however, that no amount of training will produce natural-sounding speech in the case of a severely deaf child, the deficiency being due mainly to incorrect control over the pitch of the voice. It has been shown[1] that pitch changes constitute an important factor in the intelligibility of speech. Another important factor is correct rhythm.[2] While it seems reasonable to expect that a deaf child can glean information with respect to rhythm by using his senses of touch and vision only, it is clear that the concept of pitch is as incomprehensible and abstract to such a person as is the concept of color or hue to someone who was born blind.

It was, therefore, concluded that the need existed for introducing information about voice pitch, and preferably also voice rhythm, to the brain via one of the other senses in a more effective way than by simply endeavoring to feel these effects by means of the fingers laid upon the speech organs of the teacher. It is well known that the tactile sense can detect changes in the intensity of a vibration but is unreliable and insensitive when it comes to changes in the vibrational frequency.[3] Of the remaining senses it seems logical to employ sight for observing this information, as the visual sense is already trained to accept a large amount of information and has the ability of observing at least two (and possibly three) dimensions simultaneously. In the instrument to be described, therefore, pitch is displayed along a vertical axis against a continuous horizontal time base, and, as the screen on which this graph is traced out has image storage properties, a "memory" is in effect provided in the display. As the eye, therefore, scans the information, it can recheck portions which were missed or incorrectly understood during the first "reading."

Throughout this paper the word *pitch* should be understood in its objective sense, meaning the frequency of the fundamental component of the complex voiced sound. Of course, not only the pitch is displayed, but also the duration and rhythmic pattern of all voiced sounds.

The necessity of describing such an instrument could be questioned on the grounds that a visible speech translator can provide the same and more information about speech. Such instruments, as well as associated pitch indicators, have been adequately described,[4–7] but it is felt that this new approach is justified on the grounds that the present apparatus has been specifically designed for use with young deaf children and hence has been made to display only the required information about pitch changes in the simplest possible form, so as to be comprehensible to such a child even at an early stage in its development. The formation of good or bad speech habits takes place at an early age, and these are not easily changed later in life. Used as a teaching aid, the function of the instrument should be easily grasped, and learning to interpret the patterns which it presents should not require an extensive training program, as it is difficult to find time in the already overburdened school syllabus for such activities.

[1] Charles H. Voelker, Am. Annals of the Deaf **80**, 3, 346–249 (1935). The results of a study of pitch changes in the speech of deaf people compared with those of people with normal hearing is presented.

[2] C. V. Hudgins and F. C. Numbers, Genetic Psychology Monographs **25**, 388–390 (1942). It is shown that sentences spoken with correct rhythm stands a 4 to 1 chance of being understood over those spoken with incorrect rhythm.

[3] V. O. Knudsen, J. Gen. Psychol. **1**, 320 (1928). Knudsen conducted an investigation into the sensitivity and sensibility characteristics of the sense of touch to a vibrating body.

[4] J. C. Steinberg and N. R. French, J. Acoust. Soc. Am. **18**, 1 (1946).

[5] W. Koenig, H. K. Dunn and L. Y. Lacy, J. Acoust. Soc. Am. **18**, 1 (1946).

[6] R. R. Reisz and L. Schott, J. Acoust. Soc. Am. **18**, 1 (1946).

[7] R. K. Potter, G. A. Kopp, and H. C. Green, *Visible Speech* (D. Van Nostrand Co. Inc., Princeton, New Jersey, 1947).

DESIGN CONSIDERATIONS

The most important requirement is that the instrument should display only the pitch of voiced sounds, without error and without requiring skilled manipulation. No errors should be introduced on account of insufficient or excess sound level, tonal quality, or noise content of the voice. If a trained ear is capable of distinguishing a definite pitch in the sound produced by the child, so should the instrument be. In practice, the teacher speaks into the microphone, saying the word or sentence and controlling his voice pitch in the correct way. He then hands the microphone to the child, who endeavors to imitate the patterns produced by the teacher. The teacher may be a man whose voice pitch falls in the range 100–150 cps. The child's voice may naturally fall into the range 300–450 cps. To render the two patterns directly comparable without having to switch over to a new range, a frequency scale covering at least the range 100 to 600 cps is required, and if the two patterns quoted are to look alike when they sound alike to the ear, the frequency scale has to be approximately logarithmic.

BACKGROUND

The remarkable ability of the ear to assign a definite pitch to any recurrent waveform, even when the fundamental component thereof is weak or, in the limit, nonexistent, is a well known phenomenon. In attempting to design an instrument which will do the same thing, severe difficulties are encountered. Conventional frequency meters actually measure half the number of zero crossings of the waveshape occurring per second of time and not the frequency of the fundamental component thereof. Hence, these cannot be used directly for the measurement of voice pitch. The seemingly obvious solution of preceding the frequency meter by a filter which passes only the fundamental is only applicable if the range within which the pitch frequency falls does not exceed one octave.

In an instrument described by Reisz and Schott[8] use is made of a filter operating on the above principle, and provision is made to select manually either of two octave ranges. Furthermore, the fundamental is enhanced by the expedient of allowing intermodulation to occur between successive harmonics of the speech signal in a nonlinear circuit element. This system cannot be used in the present case, as manual switching is to be avoided and the required range is more than two octaves in extent.

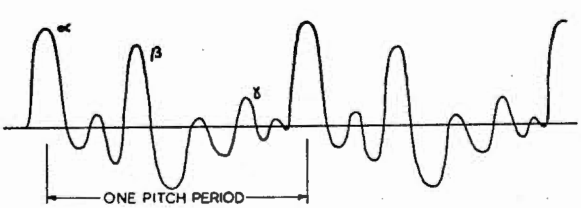


FIG. 1. A typical example of the waveshape produced by a male voice at low frequencies. Note the presence of a large amplitude swing β following the initial peak α.

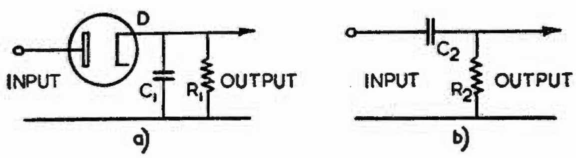


FIG. 2(a). A detector stage with storage condenser C_1 and discharge resistor R_1.

FIG. 2(b). The differentiating circuit, which is in cascade with (a), removes the lower frequency components including that at zero frequency.

The logical enquiry at this stage is whether automatic selection of the appropriate octave range cannot be achieved. An investigation of this possibility was conducted by Professor R. W. Guelke of the University of Cape Town. The range of 100–600 cps was split into three contiguous bands by means of filters, and the speech signal was applied to the input terminals of these three filters. The output of each filter was rectified and smoothed to obtain a control voltage which was used to block all filters except the one passing the lowest-frequency component. This approach proved to be ineffective when dealing with speech, as the time delay inherent in any control system gives rise to momentary incorrect operation of the blocking devices when a wavetrain is suddenly applied to the input. As speech is a continuous succession of such wavetrains, these false responses occur too frequently to be ignored.

In the Coyne Pitch Indicator, an instrument which was widely used until recently, speech signals were applied to 20 contiguous mechanical filters in the form of electrically driven tuning forks. Each filter actuated one lamp in a vertical row of 20 lamps.[9] Here again the presence of a signal from any filter was arranged to block all the filters higher in frequency than this one. Thus, on a sustained sound only one lamp was lit, corresponding to the frequency of the fundamental. The blocking action was instrumented mechanically by means of vibrating contacts actuated by the movement of the tuning fork tines, and a system of relays. It is evident that a time lag is again introduced with its attendant difficulties as outlined above. A further difficulty was that in practice the bandpass of the individual filters was insufficient, and, therefore, gaps existed in the frequency spectrum. There was also the limitation of having only a unidimensional display— there was no memory. In spite of its shortcomings, this instrument was the first of its kind and stimulated interest in the possibilities of improving the speech of the deaf by providing them with means for monitoring the pitch of their vocal efforts.

[8] See footnote reference 6, p. 57.

[9] A. E. Coyne, The Volta Review **40**, 10, 549–552 (Oct. 1938). Description and practical application of voice pitch indicator.

FINDING THE FREQUENCY OF THE FUNDAMENTAL

The present instrument in modified form uses a principle already described.[10] The problem of finding the frequency of the fundamental component of the voice is approached from an entirely different direction by investigating the mechanism by which voiced sounds are produced and then proceeding to make use of the basic characteristics of these sounds. Briefly stated, voiced sound is produced by movements of the vocal cords interrupting the smooth flow of air from the lungs, thus providing energy in the form of a series of impulses. These contain an abundancy of harmonics, some of which are reinforced by resonances in the vocal tract, while others are suppressed. The frequency spectrum of the sound produced is, therefore, under the control of the speaker by changing the shape and size of the vocal cavities, while the pitch of the sound corresponds to the pulse repetition frequency. Each pulse produces basically a damped oscillation in the vocal tract (Fig. 1), and it becomes clear that the commencement of each fundamental cycle of sound is characterized by (a) maximum slope and (b) maximum amplitude.

A circuit is sought which will recognize these two properties and produce one output pulse per fundamental pitch period. A suitable solution is shown in Fig. 2(a). The waveform of Fig. 1 is fed to the diode D, which charges C_1 as long as the input voltage exceeds that present across C_1. Beyond the peak of the first swing α conduction in D ceases and C_1 discharges exponentially through R_1 as shown by the dotted curve in Fig. 3(a). During this decay there may be further points on the waveform, such as β and γ, where the input voltage momentarily exceeds that on C_1, when D will again conduct and C_1 will receive further small charges. The final waveshape appearing at the output in Fig. 2(a) then resembles Fig. 3(b). This is then partially differentiated in the circuit of Fig. 2(b) to produce the waveshape shown in Fig. 3(c). Comparing Fig. 1 with Fig. 3(c), it is at once evident that the initial peak has been noticeably enhanced. The two basic properties of the first peak are responsible for the operation of this circuit, maximum amplitude being recognized by the diode and C_1 and maximum slope by the differentiating network. This process can be repeated as many times as is required to remove all but the initial peaks of the original waveshape. In each stage the ratio of the amplitudes of initial to subsequent peaks is increased.[11]

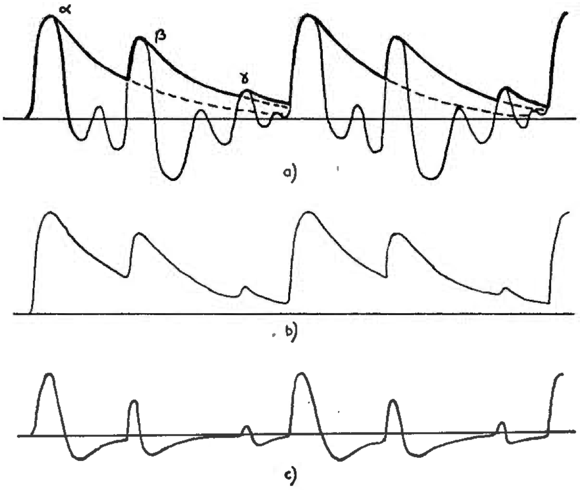


FIG. 3(a). Showing the waveshape across C_1 of Fig. 2(a). Note that the peaks α, β, and γ cause conduction. Dotted lines show the way in which C_1 would discharge in the absence of additional peaks following the first.

FIG. 3(b). Solid line of Fig. 3(a) redrawn to show the output waveshape of the circuit of Fig. 2(a) more clearly.

FIG. 3(c). Output of the circuit of Fig. 2(b). Note that the ratio of initial to subsequent peaks has increased over that of the original wave as seen in Fig. 1.

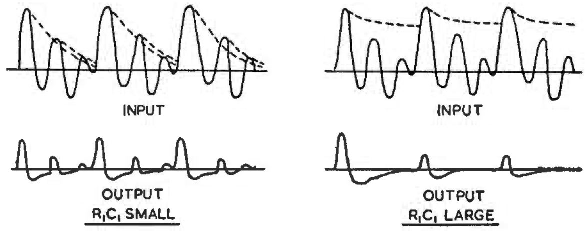


FIG. 4. Showing the effect of changing the time constant R_1C_1. When R_1C_1 is small, peaks following the first become prominent. When R_1C_1 is large, the amplitude of all pulses after that produced by the first cycle of a suddenly encountered wavetrain is attenuated.

It is of importance to choose a suitable time constant for the product C_1R_1. Making this small will allow subsequent peaks to become large, thus requiring more stages of detection and differentiation; making it large causes the output to become very small. This is shown in Fig. 4. The final value of 5 msec is a compromise which in practice is satisfactory for the frequency range concerned. For the same reason R_2C_2 was made 0.5 msec.

Once the speech wave has been simplified to the extent that there is only one pulse per pitch period, it is, of course, a simple matter to measure its frequency. The pulses will not be uniform in shape and amplitude, but they can be employed to trigger a one-shot multivibrator which shapes these into the required constant form and size, and a low-pass filter or some form of integration following this stage will produce an output proportional to the frequency.

[10] Ladislav O. Dolansky, J. Acoust. Soc. Am. **27**, 1 (1955).

[11] The speech wave is in general not symmetrical about the zero axis. The microphone is originally poled in such a way that the direction in which the highest peak occurs is that which causes current flow through the diode. As the polarity of this peak is determined by the direction in which the breath flows, and as speech is invariably produced upon exhaling, there is no need to change the polarity for different speakers. There is, however, no telling what a deaf child will do!

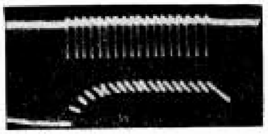


FIG. 5. Photograph of a twin-trace oscillogram showing (top) multivibrator output pulses and (bottom) the output obtained from an integrating network fed with these pulses. Note gradual rise at the beginning of the pulse train and the considerable ripple present on the output.

There is, however, a time lag associated with any low-pass filter or integration process which gives rise to false indications at the commencement of a suddenly encountered wavetrain as illustrated in Fig. 5, and means are sought to eliminate the characteristic slow rise to the ultimate value, as otherwise the response to a speech signal will contain many artifacts. A possible solution to the problem lies in measuring not the frequency of the pulses produced by the multivibrator but rather the period between successive pulses. It should then be possible to obtain a signal corresponding to the time interval between the first and second pulses of a suddenly encountered series upon receipt of the second pulse of such a series. The inherent time delay is then one pitch period only, after which the correct output is obtained immediately. Furthermore, the output should remain at this level until the third pulse is received and should then take on a new value corresponding to the time interval between the second and third pulses, and so forth.

PITCH PERIOD MEASUREMENT

The principle of operation of a suitable system for measuring this period is explained with reference to Fig. 6. The 40-μsec. pulses emerging from the multivibrator are used to close the switches S_1 and S_2, the latter via a circuit which delays the pulse by about 50 μsec. The first pulse of a series, therefore, connects C_4 to C_3, which is at that time entirely discharged by virtue of the presence of R_3 across it. The capacity of C_4 is only about 1/50 of that of C_3, and C_4 therefore always assumes a potential very nearly equal to that existing across C_3 at the instant S_1 reopens. About 50-μsec later S_2 closes, allowing C_3 to charge to the potential V. At the trailing edge of the delayed pulse S_2 reopens and C_3 commences to discharge exponentially through R_3. On the leading edge of the second pulse S_1 closes paralleling

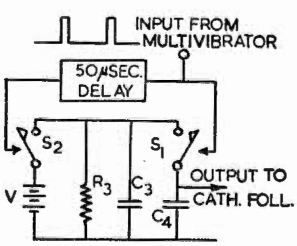


FIG. 6. Simplified diagram of period-measuring device.

C_4 and C_3. When S_1 reopens at the trailing edge of the pulse, C_4 is left floating, charged to the potential present across C_3 at that instant. This potential then, is a function of the time constant R_3C_3 and the time elapsed between the beginning of the discharge of C_3 and the instant at which S_1 reopens if $C_4 \ll C_3$. This time is about 50-μsec less than the pitch period—a discrepancy of not more than 3% which is negligible in practice.

By suitably choosing the time constant R_3C_3, the potential retained by C_4 can be made approximately proportional to the logarithm of the pitch period and by minimizing leakage of charge from C_4 by using a cathode-follower stage after C_4, this potential will stay sensibly constant over one pitch period after which it will be corrected upon the arrival of a new pulse to correspond always to the duration of the preceding pitch period.

Oscillographic records of the operation of the above circuit are shown in Figs. 7 and 8. The practical circuit used employs electronic switches for S_1 and S_2 (Fig. 6) and a regulated voltage instead of the battery.

REJECTING UNVOICED SOUNDS

The output of the pitch-measuring circuits should be zero for anything but voiced sound and means are

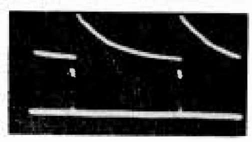


FIG. 7. Bottom: input pulses to the circuit shown in Fig. 6. Top: exponential decay of potential across C_3 showing that "fly-back" occurs a short time after the input pulse.

sought for establishing this condition. Sound spectrograms provide ample proof that in voiced sounds the energy distribution falls with increasing frequency, while in unvoiced sound, such as the fricatives, the opposite is true.[12] If, therefore, the composite speech waves are passed to two filters (a) low-pass with cutoff frequency at, say, 800 cps to pass the major portion of the energy in voiced sound and (b) high-pass with cutoff frequency at, say, 4000 cps to pass unvoiced sound energy in a portion of the frequency spectrum in which very little voiced sound energy falls, the signals from these filters rectified separately and thereafter subtracted from each other, a unidirectional output will be obtained whose sense will depend upon whether a voiced or an unvoiced sound was being presented to the input.[13] After removal of the ripple on this signal in a suitable low-pass filter (cut-off frequency about 50 cps) it can be used to block or activate the signal path in the pitch determining circuits. In the case of the present instrument it is also applied to the pitch display

[12] G. A. Kopp and H. C. Green, J. Acoust. Soc. Am. **18**, 1 (1946). Numerous examples of sound spectrograms are given.
[13] See also footnote references 6, p. 59, and 9, p. 491.

unit, as will be described subsequently. Figure 9 shows oscillographic records of the various waveshapes present in the instrument when the word "suit" is spoken into the microphone.

DISPLAY UNIT

It has been mentioned that pitch changes are to be displayed on a continuous time base and that the patterns produced should be visible for some time. The need for a continuous time base, as opposed to the customary repetitive type employed in a cathode-ray oscilloscope, arises from practical considerations. It would not be feasible to expect a young deaf child to understand that it has to wait until the time base is ready to start its sweep (if it is self triggering). If the time base sweep is arranged to be triggered by the commencement of speech, it will also be set off by background noise, the child clearing its throat before speaking, etc. Reflection on the situation indicates that a continuous time base is the only feasible solution in this case.

In visible speech translators, use is made of a special cathode-ray tube[14] or a horizontally moving phos-

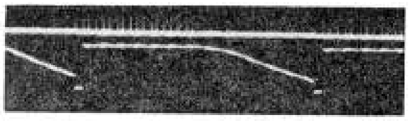


FIG. 8. Top: multivibrator output pulses obtained from a sudden input wavetrain; Bottom: output of period-measuring device showing that the first pulse encountered resets the output to zero followed by a very fast rise (which does not even show on the photograph) to the correct level. Note very slight ripple and slow decay at end of pulse train. The latter part is not seen on the display unit as the trace is extinguished when sound ceases. (Uneven spacing of pulses is due to mechanical defects in the oscilloscope camera used. A trace of hum in the intensity is also evident.)

phorescent belt exited by a vertical row of miniature incandescent lamps.[15] As neither of these was available to the author, a solution to the problem was sought which would not involve the use of highly specialized and expensive equipment. This was found in the form of a normal type of magnetically focused and deflected cathode-ray tube with a long-persistence screen. The following scheme was adopted:

The 9-in. tube is arranged to rotate, by means of a small motor, counterclockwise about a horizontal axis passing through the center of the screen and at right angles to it, at the rate of about 2 rpm. The tube is held in position by means of three rollers resting on the glass just behind the screen and three more situated around the base of the tube and resting on the slip-ring assembly as shown in Fig. 10. Motion is transmitted

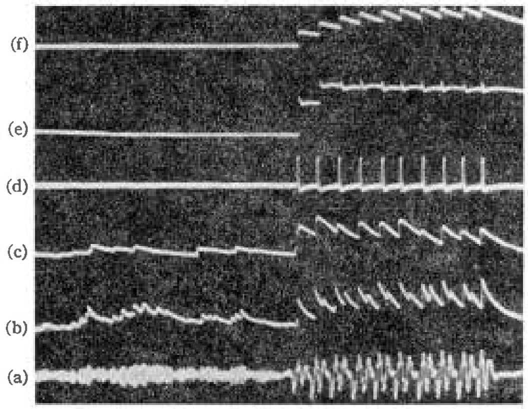


FIG. 9(a). Oscillogram of the first part of the word "suit" as spoken by a male. The *t* is not shown. Pitch: 120 cps. Note that vowel lasts for only 10 complete cycles, (b) Waveshape across C_1 of first shaper stage. For the purpose of this illustration, the microphone polarity has been reversed. The positive peaks of (a) thus produce (b) etc. Note the presence of secondary peaks. (c): as (b) but for final shaper stage. Secondary peaks have now disappeared. (d) Multivibrator output. Note suppression of unvoiced part of the word (*s*). (e) Output of period measuring device. First step is to zero level (electronic switch contact potential causes resting value to drop below zero.) Second step is to correct value representing period between second and third pulses. The spikes are due to switching transients which are readily removed by filtering without introducing excessive time lag, as these are of relatively short duration. (f) Slowly rising response and large ripple obtained if simple integration of (d) is attempted.

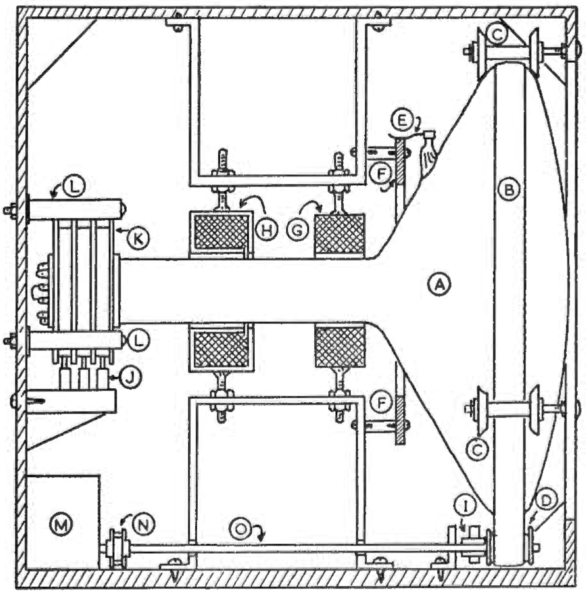


FIG. 10. Display unit. A: cathode-ray tube; B: leather driving belt; C: three equally spaced rollers supporting the front of the tube. D: driving pulley; E: brush rubbing on slip-ring F: E.H.T. connection; G: deflection coil; H: focus coil; I: front bearing and belt tensioning device of drive mechanism; J: three brushes rubbing on three slip rings K, clamped to base end of tube for conducting heater current, cathode return and control grid connection (cathode-ray tube has triode gun); L: three rollers, equally spaced around slip-ring assembly to support the base end of the tube; M: driving motor mounted as far back as possible to reduce effects of stray magnetic field on electron beam; N: flexible coupling between motor shaft and driving shaft O.

[14] J. Appl. Phys. **17**, 91 (1946).

[15] Homer Dudley and Otto O. Gruenz Jr., J. Acoust. Soc. Am. **18**, 1 (1946). In correspondence with the Bell Telephone Laboratories Inc., the present author was informed that ". . . the grain-of-wheat lamp would hardly be satisfactory because of its very short life," and again ". . . storage on an external phospor belt would be desirable, but is probably not obtainable."

from the motor to the tube by means of a flat leather belt. The required electrical connections to heater, cathode, grid, and anode are made via four slip-rings. The focus and deflecting coils are stationary and accurately mounted coaxially, so that adjustments of focus and spot position will not suffer when the tube rotates. Only one deflecting coil is used, and is positioned in such a way that the radial deflection is at an angle of about 30° to the vertical, as shown in Fig. 11. The no-signal position of the spot is at L, and the point H corresponds to the spot position when a signal of the highest pitch is being displayed. As the tube rotates, the afterglow of the spot progresses from right to left. The beginning of an utterance is, therefore, seen on the left-hand side of the pattern, followed by subsequent events to the right. In scanning the pattern, the eye moves from left to right, as in the reading of print.

Signals from the circuit which distinguishes between voiced and unvoiced sounds are applied to the cathode-ray tube grid in such a way that the electron beam is cut off on unvoiced sounds and during silent periods. This gives further protection against the display of undesirable signals. In practice it was found advantageous to give some indication of the presence of an unvoiced sound. This was done by arranging for a small neon lamp to glow when the output of the above circuit indicates the presence of an unvoiced sound.

Except for a sector at the top, the entire screen surface of the cathode-ray tube is covered by a mask (see Fig. 11). The fluorescent spot itself is situated just

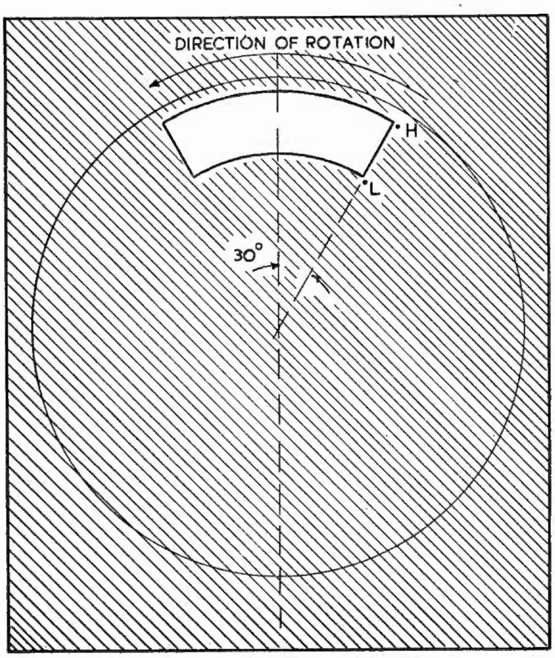


FIG. 11. Front view of display unit, showing mask which covers the face of the tube and leaving only a sector at the top exposed for viewing the trace. The resting position of the electron beam is at L, and it moves over the range L to H, representing pitch frequencies of 100–600 cps.

beyond the right-hand edge of this sector and therefore cannot be seen directly. However, as the tube rotates the afterglow appears from behind the mask within a fraction of a second after being "written." There is, of course, a steady decay in the afterglow, but at the speed of rotation which allows the width of the sector to represent about 5 sec in time the brightness of the entire pattern visible in the sector is sufficient in practice to permit comfortable viewing under subdued room-lighting conditions. By the time the tube has made a complete revolution, however, the patterns have decayed below visibility and the phosphor is ready to receive a fresh pattern.

CIRCUIT

A block diagram of the circuit used is shown in Fig. 12, together with typical waveshapes present at the various stages. Components which have not been described will be dealt with now.

Detector Stages C, E, and G

These have to be directly coupled to fairly low-impedance drivers. A source resistance not exceeding about 15 000 ohms is satisfactory when using a 0.02 MF condenser for C_1, as in Fig. 2. Had condenser coupling been employed, such a condenser would share the available charge with C_1, which is of course undesirable, as after a number of cycles of the input wave form C_1 would not receive any further charge. Direct coupling requires that the return circuit of the diode stages be brought to a point having the same dc potential as the resting potential of the source. This could have been done by using transformer coupling, but a more economic method is shown in Fig. 13. Here the return path is automatically held at the same potential as the average dc potential of the supply point by virtue of the large time constant of C_5R_4. Slow drifts in the supply point potential are followed, but even the lowest signal frequency fails to affect this potential significantly.

Isolating Amplifiers D and F

These are inserted between stages C and E and again between E and G of the waveform shaping circuits in order to make up for losses in amplitude as well as for isolating and impedance matching purposes.

Adding ~~Added~~ Circuit S

It was mentioned before that the rectified output signals of the two channels in which voiced and unvoiced sounds are separated from one another are subtracted. This is accomplished by addition in an elementary resistive network, after first reversing the polarity of one of the rectifiers R (that in the unvoiced channel). Hence the polarity of this control signal is as shown in Fig. 12. Full-wave bridge rectifiers are used to facilitate smoothing.

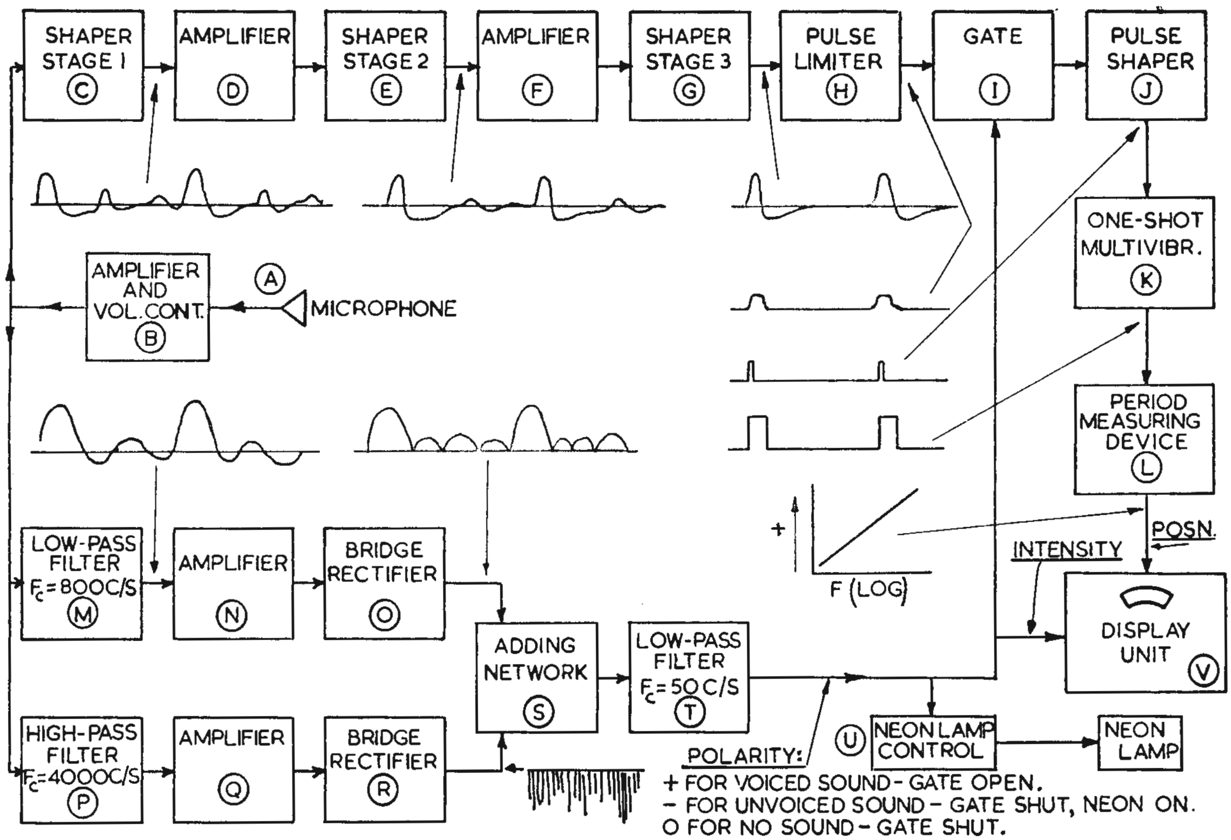


Fig. 12. Block diagram of pitch indicator with typical waveforms at several points in the circuit.

Gate *I* and Pulse Limiter *H*

A pentode operating from a low anode-supply voltage is used in the gate stage. The pulse signal from (*H*) is applied to g_1 of this tube and the control signal to g_2 which draws only a fraction of one ma. The circuits *S* and *T* can supply this small current quite readily. When the control signal is zero or negative, this tube does not conduct and effectively blocks the signal path. When positive, it not only conducts but also assists in limiting the pulse amplitude in conjunction with the previous stage *H*. For reliable operation of the one-shot multivibrator, it is essential that the triggering pulses fed to it should not vary too widely in shape and size. As the dynamic range of the sound falling on the microphone can be quite large, and further expansion of this could occur in the waveform shaping circuits, it is important that pulse limiting should be effective.

Filters

As no stringent requirements need be placed on any of the filters *M*, *P*, or *T*, these are not of the L-C type. It was convenient and advantageous (as they also provide the required gain) to use active R-C filters for *M* and *P*[16,17] and a simple cascaded R-C filter for *T*.

[16] C. H. Miller, Wireless Engineer **27**, 316 (1950).
[17] A. R. Thiele, Electronic Eng. **28**, 335, 31–36; and **28**, 336 (1956).

Deflection Amplifier

This forms part of the display unit and provides drive current for the deflection coil of the cathode-ray tube. The resting anode current is set to a suitable value to deflect the spot to the lower edge of the aperture in the mask. On this are then superimposed the signals of the period measuring device *L* causing further radial deflection of the spot. The relationship of the spot position to the pitch of the input signal is approximately logarithmic.

EHT Generator

A supply of 4000 v is required for the final anode of the cathode-ray tube. This is obtained from a series-

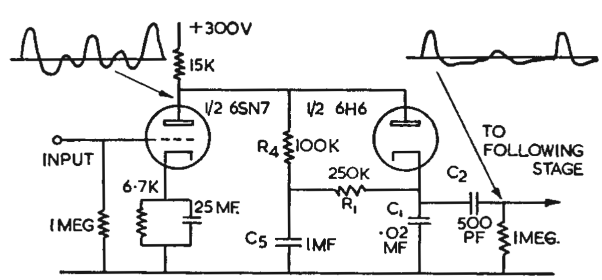


Fig. 13. Practical diagram of one stage of shaping circuits, showing direct coupling between driver stage and diode, with input and output waveshapes.

tuned, radio-frequency type of generator operating at approximately 280 kc/sec. The high voltage is rectified by a diode whose filament is heated by the circulating current of the tuned circuit.[18]

PERFORMANCE TESTS ON APPARATUS

With no sound falling on the microphone, the bias on the control grid of the cathode-ray tube is adjusted to just extinguish the fluorescent spot. As a good signal-to-noise ratio is required, it is good practice to hold the microphone within a few inches of the lips and to use barely sufficient gain to ensure reliable operation, unless the instrument is used in a quiet, acoustically treated room. Keeping the voice pitch steady, it is possible to produce any vowel sound or nasal consonant and get exactly the same indication of pitch on all. The instrument is also quite insensitive to such changes in voice quality as that which exists between different speakers. When the voice pitch is changed smoothly or in discrete steps, the display follows the pattern reasonably well. Figure 14 shows the display obtained when singing the major scale into the microphone. Note that no false indications are present and that only the voiced parts of the words are displayed. Note also that the untrained voice used when taking this photograph glided up to correct pitch on the sixth and eighth and produced the fourth with a marked vibrato. The only sounds on which false indications are occasionally produced (the display jumping an octave) are the *u* as in "foot" and the *i* as in "seek." These sounds contain very prominent peaks following the initial one. Such false indications are, however, rare and only occur in the case of very low-pitched male voices.

The circuits which determine whether a sound is voiced or unvoiced perform satisfactorily, except in the case of some of the plosive stop consonants like "ste*p*," which do not contain strong high-frequency components but rather appear in the form of a single pulse made up of a restricted band of low-frequency (mainly below 100 cps) components. They, therefore, activate the gate circuit, brighten the display, and are shown in an erratic manner just below the 100 cps position on the screen. The opposite effect (voiced sound, not displayed) has been known to take place in the case of an exceptionally piercing male voice producing the sound *i* as in "seek" where abnormally strong harmonics are present within the range of the filter, which selects high frequencies in order to render the display inoperative.

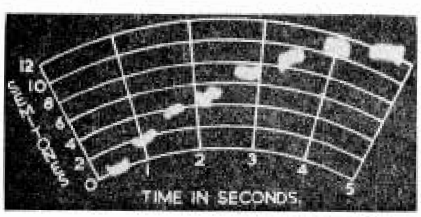


FIG. 14. Display unit presentation of the notes of the major scale sung by a male voice pronouncing the words: doh, ray, me, etc., with a time and frequency scale superimposed photographically to show closely logarythmic response. The frequency scale was contracted for the purpose of this illustration from the normal $2\frac{1}{2}$ octaves to 1 octave.

[18] Peter G. Sulzer, Electronics **25**, 9 (1952).

The neon lamp which glows on unvoiced sound performs very satisfactorily. Quite unexpectedly to the author (who does not claim to be a phonetician), it has proved in practice to be a valuable adjunct, showing as it does the existence and duration of a breath sound at those times when no pattern is being formed on the screen. The existence of fricatives, stop consonants, etc., is therefore clearly indicated to the deaf child. As these parts of speech are hard to see in lipreading, all deaf children find them difficult, and even limited assistance in this connection, such as that rendered by this neon lamp, is of value.

DESIRABLE IMPROVEMENTS

Brighter Display

Although it has been stated that the afterglow is visible over the entire width of the aperture in the mask under subdued lighting conditions, it would be desirable to get a brighter display as it is not always convenient to get such conditions in practice and also, in dealing with deaf children who have to lipread their instructions, a higher level of ambient lighting is an advantage. Achieving the desired brightness of display is, of course, merely a matter of using a tube to which a higher-accelerating potential can be applied.

Longer Time-Persistence

This can safely be increased, as in the case of the cathode-ray tube used at present the images have decayed to a level below visibility long before the tube has made one revolution when moving at an angular velocity which has proved to be adequate in practice (2 rpm). It would be better still if the screen were of the image-storage type in order that the patterns may be viewed for any desired length of time (by stopping rotation after a pattern has been "written" onto the screen) and subsequently erased automatically just before that part of the screen was required afresh. If this could be achieved, the teacher could point out to the child at leisure and in detail just where it went astray.

Improved Contrast Range

In the present instrument, owing to the low trace brilliance, no intensity modulation could be applied to the cathode-ray tube (other than the on-off modulation). In an improved tube this is a possibility worth considering, as it will then be possible to present a third dimension, such as voice intensity.

PRACTICAL USE AND RESULTS

The present experimental apparatus has been in use for some considerable time at this school and has enjoyed interest among similar institutions elsewhere. In particular has it been used by a group of eight children whose ages range from 8 to 12 years and whose hearing loss is from 60 db upwards. None of them can understand speech by hearing only.

It was found that the concept of pitch is so abstract to a young deaf child that at first the patterns presented by the pitch indicator are quite meaningless to him. In training the children to use the instrument, it was therefore found necessary to proceed along the following lines:

1. Learning to observe that the duration of a voiced sound corresponds to the horizontal length of the trace produced. This is useful for teaching the difference between:

(a) long and short vowels:
ar (cart) and *a* (cat)
oo (moon) and *oo* (book)
ee (sheep) and *i* (fish);

(b) words in which such vowels are found:
a c*ar*t and a c*a*t
a m*oo*n and a b*oo*k
a sh*ee*p and a f*i*sh

(c) phrases involving words containing long and short vowels:
a bl*ue* c*ar*t . . . a bl*a*ck c*a*t
a b*i*g f*i*sh . . . a g*oo*d sh*ee*p.

2. Learning to observe that a voiced sound produces a pattern on the screen and no red light, while an unvoiced (breath) sound produces a glow of the light and no pattern. This is useful for teaching:

(a) the difference between long voiced consonants like m, n, ng, v, z, th (then), r- (run), y- (yes) and breath consonants like p, t, k or c, ch, wh, f, s, th (thin), sh;

(b) correct speech in cases where pronunciation is slovenly:
I. where the *s* or *ed* or *t* is dropped at the end of a word, e.g. duck(s), sum(s), walk(ed), sal(t);
II. where the initial *s* is dropped, e.g., (s)treet, (s)cream, (s)chool;

(c) that the consonants m, b, and p look alike in lipreading, but are vastly different from one another when pronounced. It is a very common weakness among the deaf to substitute voiced for breath consonants and vice versa, because to them all these parts of speech look alike. They may substitute: man for pan, pump for bump, etc. Confusion often also arises in the case of the gutteral stop consonants c and g (cold and gold) and in the n, t, d group. They might say non or nont or dod for don't and ninin for didn't. In all such cases differentiation is now possible by observing the screen and neon lamp, for one sound is voiced and the alternative sound is unvoiced.

FIG. 15. As Fig. 14, but showing the pitch patterns produced by saying the words "good morning" in two different ways. Accented syllables are underlined.

3. Learning that speech has rhythm:

(a) accent in words: wá/tĕr, tō/dáy, tō/mór/rōw.

(b) accent in phrases: ā cát, ā bút/tĕr/fly, in̄ t̄he hoúse, gōōd mór/nin̄g, nó thānk yōu.

(c) rhythm in counting: ON̄E twō thrēe, ōne TWO thrēe, ōne twō THRÉE.

(d) to show that the meaning of a sentence can change according to which word is stressed:
The *man* went to the shop yesterday. (who)
The man *went* to the shop yesterday. (affirmative)
The man went to the *shop* yesterday. (where)
The man went to the shop *yesterday.* (when)

It is, of course, commonly found that when a word is stressed the voice pitch also rises on that word. This serves as an introduction, therefore, to the understanding of pitch changes. An illustration of this is given in Figs. 15 and 16.

4. Learning the meaning of voice pitch. Once this is grasped, the child is ready to receive instruction to correct the way of changing his voice pitch, e.g.:

(a) The voice drops at the end of a sentence.
(b) The voice rises at the end of a question.
(c) Proper pitch control is required to make exclamations more interesting and natural.

The voices of some children are very high or very low in pitch. There often is no change of pitch at all when they speak, or changes take place in a most unnatural way. Many of them speak very slowly and prolong the vowels unnecessarily. Using the pitch indicator, these discrepancies can now be demonstrated in a form that

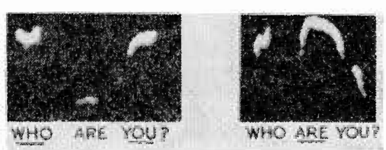


FIG. 16. As Fig. 15, but for the phrase: "who are you."

the child can grasp, and these no longer belong, as far as he is concerned, to the realm of the abstract.

CONCLUSION

As already mentioned the present instrument has some distinct shortcomings, and, therefore, its potentialities have not been realized fully. It has, however, made possible the aquisition of experience leading to the development of a new design, which it is hoped, will be more nearly ideal for the purpose. The important fact at present is that the instrument is proving to be of practical assistance in teaching the deaf, and the children are interested in using it.

ACKNOWLEDGMENTS

The author wishes to thank Professor R. W. Guelke of the University of Cape Town, whose interest and help with this project have been an encouragement at all times, and Miss E. J. Dommisse, member of the staff of this school, who has made extensive use of the instrument and offered many helpful suggestions in connection with the practical application thereof.