

May 27, 1941.

H. W. DUDLEY

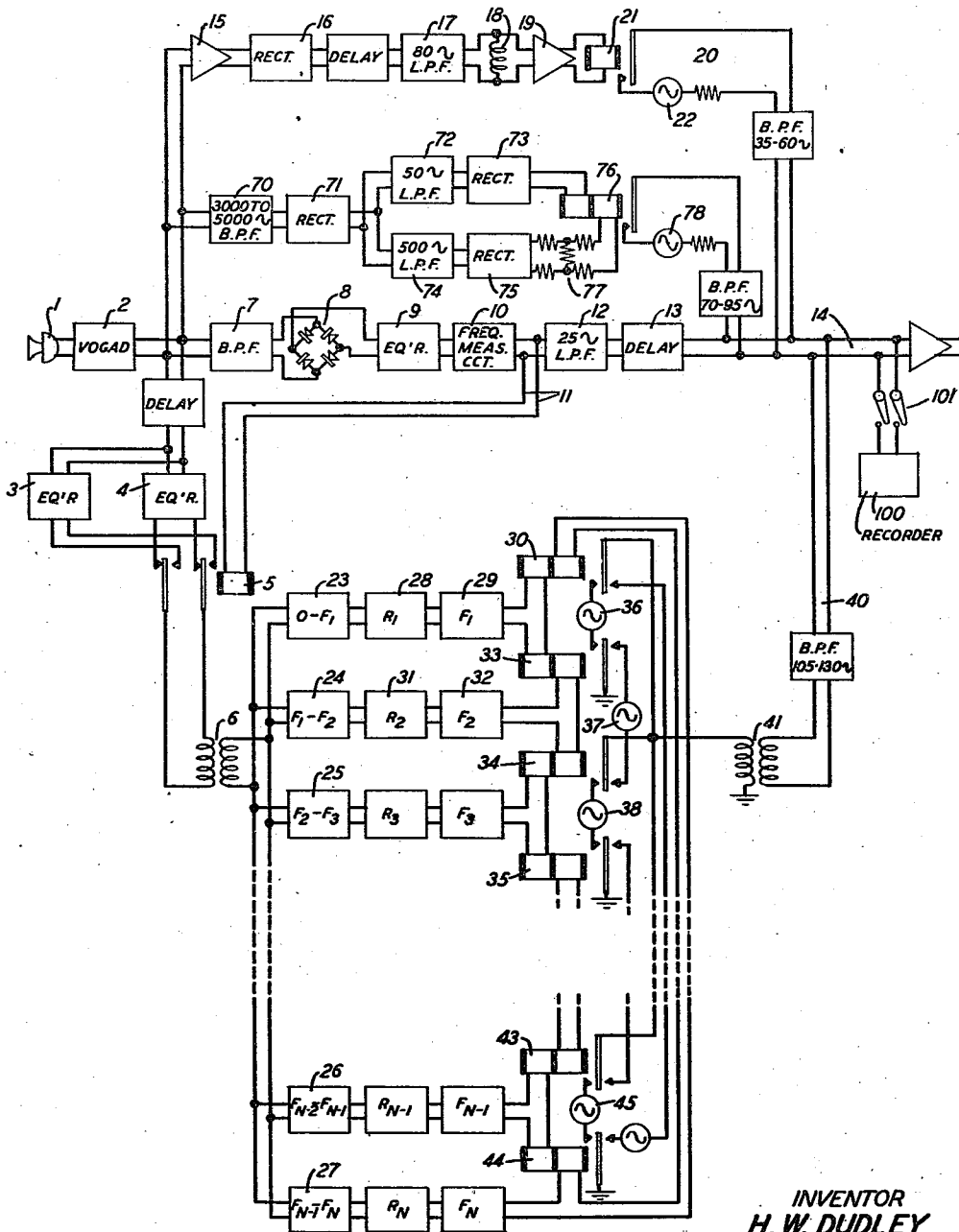
2,243,527

PRODUCTION OF ARTIFICIAL SPEECH

Filed March 16, 1940

2 Sheets-Sheet 1

FIG. 1



INVENTOR
H. W. DUDLEY

BY

H. A. Buzess

ATTORNEY

May 27, 1941.

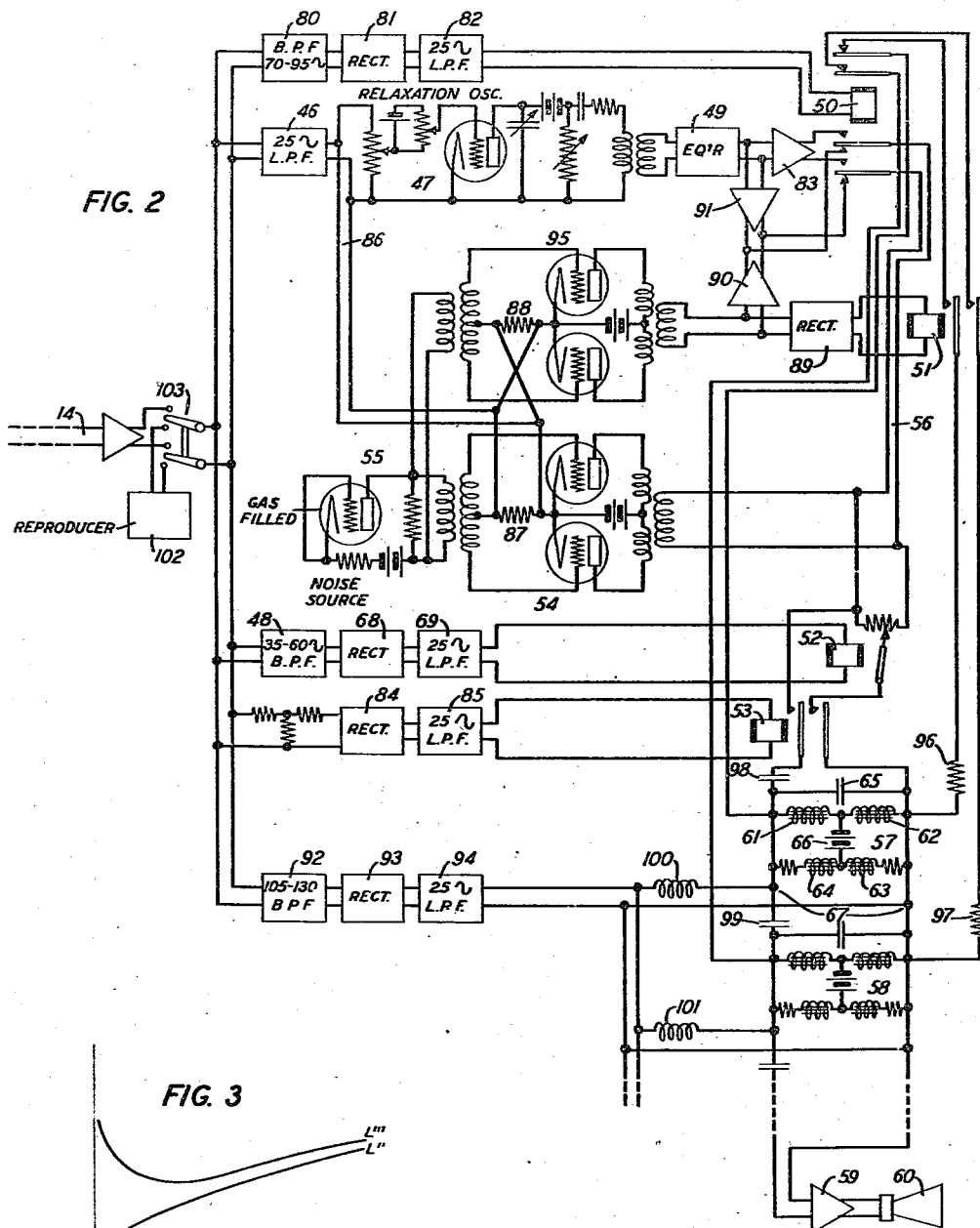
H. W. DUDLEY

2,243,527

PRODUCTION OF ARTIFICIAL SPEECH

Filed March 16, 1940

2 Sheets-Sheet 2



INVENTOR
H. W. DUDLEY
BY *H. A. BURGERS*
ATTORNEY

UNITED STATES PATENT OFFICE

2,243,527

PRODUCTION OF ARTIFICIAL SPEECH

Homer W. Dudley, Garden City, N. Y., assignor to
Bell Telephone Laboratories, Incorporated, New
York, N. Y., a corporation of New York

Application March 16, 1940, Serial No. 324,288

10 Claims. (Cl. 179—1)

The present invention relates to the artificial production of speech or similar complex waves. The invention also relates to the transmission of speech with reduced frequency range from one point where the input speech is analyzed to a distant point where the speech is reconstructed under control of the analyzer. A system of this general type is disclosed and claimed in my prior Patent 2,151,091, patented March 21, 1939. The present invention makes use of a different type of analyzer and a different type of synthesizer from those shown in my prior patent.

An object of the present invention is to simulate to an increased extent the human speech mechanism in the synthesizer and to provide a type of analysis suitable for controlling such synthesizer. In my prior patent referred to, an input speech wave is analyzed to determine its fundamental frequency and its frequency amplitude distribution or amplitude pattern over the frequency range of the speech. The result of this analysis is translated into a varying current representing the pitch and a number of varying currents in different circuits, for example, ten such currents, representing the instantaneous speech power in a corresponding number of regions of the speech frequency range. The information contained in all of these varying currents is sent to the synthesizer and used to build up from sources of energy in the synthesizer an artificial speech wave with the characteristic pitch and frequency amplitude distribution of the impressed speech.

According to the present invention, the human speech producing mechanism is more closely approached in the synthesizer by using in place of the large number of fixed filters, disclosed in my prior patent, a small number of resonant circuits, such as two, which are variable and which can be variably controlled to produce sounds that merge one into the next as the transition is made from one sound to another. The analyzer in accordance with the present invention has its construction and design determined largely by the requirements of the synthesizer and instead of analyzing the speech into a large number of varying currents as in my prior patent, it operates to give an indication of the portion of the speech band carrying maximum power at any instant. The synthesizer to be disclosed herein is similar, in general, to that disclosed and claimed in my copending application, Serial No. 324,286, filed March 16, 1940. The analyzer to be disclosed herein is a modification of the analyzer disclosed and claimed in my copending application, Serial No. 324,287, filed March 16, 1940.

The human speech mechanism makes use of the vocal cord energy in producing voiced sounds and makes use of the breath without vibration of the vocal cord energy in producing voiced sounds and modifies the effects of the breath by the movement of the tongue, lips and teeth as is well known. The voiced sounds differ from each other depending upon the action which the mouth has upon the passage of the vocal cord energy through the mouth. The modifications of the vocal cord energy produced by the mouth consist largely in modifying the resonance. The mouth cavity, for example, is divided principally into two portions by movement of the tongue and the relative sizes of these two portions vary and modify the resonances. At any instant the two air chambers may differ considerably in size and, therefore, in resonance, so that many of the vowel sounds are produced by double resonance effect, such as to emphasize low frequencies in a certain range and at the same time high frequencies in a certain range. These resonances, moreover, are not independent of each other but are limited because of the fact that the total air space in the mouth remains more or less constant.

Some of the consonants are produced with a more or less open mouth position and, therefore, have the same general characteristics as the vowels. Others are produced with the formation of decided constrictions in the mouth chamber. It is convenient to classify the speech sounds in two ways.

1. As to spectrum, whether periodic, non-periodic or a mixed type corresponding to voiced sounds, constriction type sounds and combination type sounds.

2. As to rate of formation, whether explosive or sustainable.

On this basis the sounds will be listed, the consonant sounds being mentioned specifically and the vowel sounds being treated as a group, the individuals of which have been mentioned in other of my applications. The groups are:

1. Voiced explosives—*b, d, g* (hard).
2. Unvoiced explosives—*p, t, k*.
3. Voiced sustainable sounds consisting of the vowels together with the transitionals—*w, y*, and the semivowels—*l, m, n, ng* and *r*.
4. Unvoiced sustainable sounds which are *h, wh* and the fricatives—*f, s, th* of "thin" and *sh*.
5. The voiced or mixed type fricatives, *v, z, zh*, and *th* as in "then."

In speech analyzing, some arrangement must be provided for handling the diphthongs. The assumption here is that the early part of the

diphthongs suitable for treatment in terms of the diphthong and the latter part sufficiently like the second vowel of the diphthong as to make diphthongs suitable for treatment in terms of component sounds. It may be necessary, however, to allow for certain gliding effects where resonance conditions are produced that are not in exact correspondence with the component vowel sounds. If so, the only effect is to give a few extra sounds which must be provided for and can be, by the method used for the other sounds. In the same way it may be desirable to provide for the gliding or elision in ordinary talking where the precise vocal part positions are not always obtained by a talker before he proceeds to the next sound.

The largest class of sounds is seen to be that in group 3. Fortunately, this is oftentimes the less important set from the standpoint of analysis. Thus Alexander Graham Bell pointed out that one can take a paragraph and use the same general vowel sounds for each vowel and still have the listener understand mostly what was said. In other words, errors in the vowel group are less serious from the standpoint of intelligibility than elsewhere. They are also undoubtedly less noticeable because many people pronounce them rather poorly, oftentimes substituting another vowel sound.

In accordance with the present invention to be disclosed more fully hereinafter, an electrical resonance synthesizer is provided which corresponds functionally to the air chambers of the mouth and must be able to vary at the same rates at which the portions of the mouth move in producing speech. This, incidentally, is a relatively low rate, of the order of seven per second. This means that very low frequencies suffice for controlling the synthesizer and that therefore only these low frequencies need to pass over the line from the analyzer to the synthesizer. The analyzer produces indications of three general types, (1) the fundamental frequency together with its variations with time, (2) the resonance characteristic and (3) explosive sounds.

The various objects and features of the invention will appear more clearly from the following detailed description in connection with the accompanying drawings in which:

Fig. 1 is a schematic circuit diagram of the transmitting terminal of a system according to the invention;

Fig. 2 is a similar diagram of the receiving terminal of the same system; and

Fig. 3 shows curves of certain inductance characteristics to be referred to in the description.

Referring first to Fig. 1, the wave input source for speech or other sounds to be analyzed is indicated at 1 and this may be a microphone or any other type of energy converter for converting from sound vibrations, mechanical vibrations or light variations into electrical variations. The pick-up 1 feeds into a "Vogad" (voice operated gain adjusting device) for reducing to a constant level, in its output, waves of various levels applied to its input and coming from different talkers or input sources of different energy levels. This vogad may be of the type disclosed in Mitchell-Shott Patent 2,019,577, issued November 5, 1935, or in the Hogg et al. Patent 1,853,974, issued April 12, 1932, referred to therein. As a result of using the vogad, the speech waves impressed on the analyzer are maintained at constant average level even though the talker volume at the pick-up 1 may vary widely.

The speech waves on the output side of the vogad pass into four separate branches. The uppermost branch leading through amplifier 15 is the analyzer for the explosive consonants, to be described later. The branch leading through band-pass filter 7 is the foundation frequency analyzer branch or fundamental frequency control branch, to be described later. The lowermost branch leading through equalizer 3 or 4 is the branch leading to the spectrum analyzer for determining the speech frequency subbands having predominant power at a particular instant. Equalizer 3 is used if the relay 5 is not energized (in a manner to be described), which means that the waves present are unvoiced waves. If the waves contain vocal cord energy, equalizer 4 is substituted for equalizer 3 by the fact that relay 5 is energized. The purpose of these equalizers will be explained at a later point.

The fundamental frequency control circuit extends from the output of the Vogad 2 through band-pass filter 7 whose pass range may be such as to pass the essential or important speech frequencies, for example 50 to 3,000 cycles per second. This filter selects the fundamental voice frequency component and a few of the lower harmonics thereof which are rectified in the rectifier bridge 8 to derive the fundamental frequency component of the speech. This is passed through equalizer 9 which, as disclosed in R. R. Riesz Patent 2,183,248, has its loss increasing with frequency so as to insure that the fundamental frequency, which may vary, for example, from about 80 to 400 cycles, comes out at a high power level compared to any upper harmonics that may be present. The equalizer 9 is connected to frequency measuring circuit 10 which may be of the type shown in Fig. 2 of the Riesz patent referred to and the function of which is to produce a direct current, the strength of which varies in proportion to the fundamental frequency of the input speech. This is sent through the 25-cycle low-pass filter 12 and delay circuit 13 to the circuit 14 which leads to the relaxation oscillator 47 of Fig. 2. Variations in the strength of the current transmitted over circuit 14 produce variations in the fundamental frequency of the waves produced in the relaxation oscillator 47 so that these waves follow the frequency variations of the vocal cord waves of the talker. The delay circuit 13 is included to permit time for the operation of the relays in the analyzer.

When voiced sounds are present in the pitch control circuit they cause the energization of relay 5, since direct current is then impressed on the circuit 11. Relay 5 in energizing substitutes equalizer 3 for equalizer 4 in the input circuit to the analyzer. The design of equalizers 3 and 4 is dependent upon the type of filtering used in the analyzing. If the analyzing filters are of equal band width, for example, equalizer 3 is designed to compensate for the falling off of power with increase of frequency in the case of voiced sounds, with the object of more nearly equalizing the amount of energy transmitted through each equalizing filter for the same level of input speech. In this case, the equalizer 4 might be omitted since the frequency energy distribution of the unvoiced sounds is nearly uniform. On the other hand, if the band width of the analyzer filters is chosen on a logarithmic basis, the voiced sounds for normal vocal cord output would give nearly uniform power in each filter band, in which case the equalizer 3 might

be omitted but equalizer 4 would be necessary in order to equalize the normal power output of the various filters in the case of the unvoiced sounds.

The uppermost channel of Fig. 1 which is the stop consonant analyzer includes an amplifier 15 for isolating this channel from the other circuits connected across the output terminals of the Vogad 2. The output of amplifier 15 is rectified at 16 and the rectified output is passed through a low-pass filter 17 which passes the band between 0 and 80 cycles to eliminate the fundamental frequency of any voiced sound that may be present. A large inductance 18 is shunted across the output of filter 17 after which comes an amplifier 19, if desired. The design of this circuit branch is such that when a sudden drop in the energy level occurs, such as is caused by starting a stop consonant sound, a sudden pulse of energy goes through this branch and builds up a potential across inductance 18, which is amplified at 19 and operates polarized relay 21 to connect a source 22 of waves of a particular frequency, e. g. 35 cycles, to circuit 20 leading to line 14.

The circuit branch connected to band-pass filter 70 is for distinguishing between voiced sounds on the one hand and unvoiced or mixed sounds on the other. A band from 3,000 to 5,000 cycles or some similar higher frequency range is passed through this filter and then rectified at 71 to bring out a large number of difference frequencies. These difference frequencies will be in a harmonic relation if the sound is a purely periodic one. If, however, it is of the mixed energy type with the vocal cord period present but also a constriction, then there will be a more or less continuous spectrum of energy with a likelihood that the random spectrum energy will be greater than the discrete spectrum energy since the random type is more or less flat with frequency, whereas the discrete type falls off rapidly with frequency and this high frequency range has been selected to stress this difference. The energy is next passed through a 50-cycle low-pass filter 72 which should contain no energy for a pure voiced sound but should contain energy for an unvoiced sound or a mixed type sound. Similarly some of the output of the rectifier 71 is passed through a 500-cycle low-pass filter 74 to be sure to include the fundamental of the talker. For unvoiced sounds and mixed energy sounds there should be about 10 decibels more output from this filter than from the zero to 50-cycle levels. The outputs of these two filters are then applied to respective rectifiers 73 and 75 and thence differentially to a biased relay 76. Pad 77 compensates for the mentioned difference in output of the two filters in the case of unvoiced and mixed sounds. This relay remains unoperated for the unvoiced sounds and the mixed type of sounds but is operated by the voiced sounds. Thus, voiced sounds result in connecting source 78 which may have a frequency of 70 cycles, for example, to the circuit 20 and outgoing line while mixed and unvoiced sounds do not connect source 78 to the outgoing line. The effect of these operations on the synthesizer will be described later on.

The Vogad circuit 2 should operate over periods of time slightly longer than the syllabic periods of time. If the Vogad operates over shorter periods of time, it will tend to level out the rise and fall of energy corresponding to the stop consonants and thus make the explosive circuit fail to operate satisfactorily. Alternatively, the explosive circuit may be branched off from the input 1 at a

point ahead of the Vogad 2, although in this case it might be necessary to use a very sluggish volume control direct from one talker level to another. The use of the explosive analyzer channel is advantageous, since otherwise it would be necessary to analyze in the main analyzer circuit sounds of very small time interval. This would require the speeding up of the circuits of the main analyzer. The explosive analysis should determine such unvoiced sounds as *p*, *t*, *k*, as well as the voiced sounds *b*, *d* and *g*.

The waves transmitted through either equalizer 3 or 4 and repeating coil 6 are impressed on a number of relatively narrow-band filters 23 to 27, only a few of which are shown in the drawings. These filters subdivide the speech into a number of narrow bands and the output of each filter is passed through the corresponding rectifier, such as 28, 31, etc., and the rectified current is smoothed by low-pass filter 29 or 32, etc., and rendered suitable for operating relays. As many of these subdividing filters may be used as required, for example, 30 or more. In the simplest case, a single filtered band will be found to be predominant by reason of a resonance at that point. Because of this there must be at least as many bands as there are speech sounds to be recognized.

The output of one filtering branch leads to two windings, one on each of two relays such as 30 and 33 where the output of filter 29 is shown as including the left-hand winding of each of these relays. Similarly, the output of filter 32 includes the right-hand windings of relays 33 and 34. This plan is continued throughout the entire series of relays.

Each of these relays is polarized and throws its armature to the right or the left depending upon the predominance of current in a particular direction in the relay winding. Normally, the armature occupies a central position in which it makes no circuit contact. The armatures of the various relays control the connection of current sources 36, 37, etc., to outgoing circuit branch 40 through repeating coil 41. These current sources may be direct current or alternating current of any suitable type for transmission over the circuit to the synthesizer but as shown they are assumed to be alternating current sources of different current strength. These sources may all have the same frequency, e. g., 105 cycles, as illustrated, and may in effect comprise a single source of alternating current such as a machine or vacuum tube oscillator connected through resistance pads or potentiometers to the contacts of the various relays so that a different voltage is impressed on the outgoing line for each circuit closure through the relay armatures. This arrangement is preferred since it would insure in-phase addition of voltages in case a plurality of sources indicated as 36, 37, etc., were simultaneously connected to the outgoing circuit.

If the wave being analyzed contains at a particular instant more power in the band passed by filter 23, for example, than is contained in the adjacent filter bands, the relays 30 and 33 will operate to move their armatures to the left and connect source 36 to the outgoing circuit 40. If on the other hand the wave being analyzed contains maximum power in the portion of the band passed through filter 24, relays 32 and 34 operate causing these relays to move their armatures to the right thus connecting source 37 into the outgoing circuit. The direction of winding on the relays is such that the currents transmitted

through adjacent filters, such as 23 and 24, are opposed in their effect on the same relay such as 33. Under the conditions assumed above, current passed by filter 23, as already indicated, tends to attract the armatures of relays 30, 33 to the left, while current passed by filter 34 tends to move the armature of relay 33 to the right. By connecting the relays in pairs in this way, the particular current source 36, 37, etc., which is connected to the outgoing circuit, is determined by the frequency portion of the speech band carrying the maximum energy as compared to the adjacent subbands. Since, as stated, each of these current sources has a different voltage, the frequency portion of the speech band containing a maximum of energy is translated into current of particular strength impressed on the line.

It is possible that a pair of relays, such as 33 and 34, may be operated simultaneously with another pair of relays, such as 43 and 44, responding to speech currents in an entirely different portion of the speech frequency range, for any resonant region will give a maximum as compared to adjacent frequency bands. Assuming that the voltage of the sources beginning with 36 and proceeding downward in the figure gets smaller from generator to generator, the operation of the second pair of relays 43 and 44, as assumed, and the consequent connection of current source 45 to the outgoing circuit result in the supply of a relatively small increment of current from source 45 to be added to the current from the source 37. The magnitudes of the voltages of the different sources should be graduated such that when any single source is connected to the line or any two possible sources are simultaneously connected to the line the resulting current has a value representing that particular and no other setting of the analyzer. This is facilitated by the fact that it is impossible to connect simultaneously to the line two adjacent sources, such as 36 and 37. Moreover, the two resonances which are characteristic of any voiced sound are usually widely separated in the speech frequency range. This makes possible the choice of one range of values for the generators corresponding to the low frequencies, such as generators 36, 37, etc., and the choice of a different range of values for the generators, such as 45, representing the high frequency end of the scale.

The currents resulting from the analysis are transmitted over the line 14 to the synthesizer shown in Fig. 2 which may be at the same location as the circuits of Fig. 1 or may be separated at a distance therefrom, in which case the line 14 represents a long line or a channel of any suitable type, such as a radio channel or carrier channel, suitable terminal apparatus being assumed according to the type of channel. Such terminal apparatus is old and well known in the art.

The circuit of Fig. 2 includes a relaxation oscillator 47 for generating waves of buzzer-like form comprising a fundamental and a number of harmonics similar to the case of the vocal cord waves. This relaxation oscillator may be of the type disclosed in Fig. 3 of the patent to R. R. Riesz 2,183,248, December 12, 1939. As disclosed therein, the frequency of the fundamental wave is determined by the voltage impressed across the resistance 48. The output of the oscillator 47 is passed through an equalizer 49 for giving any desired relation between the amplitude of the fundamental and that of the various harmonics throughout the speech range. For ex-

ample, this equalizer may be designed to equalize the amplitudes of the fundamental and harmonics or it may have a sloping characteristic.

The circuit of Fig. 2 includes a noise source 55 shown in the form of a gas-filled tube with a suitable plate resistance and plate voltage. This noise source may be of the type disclosed in my application Serial No. 273,429, filed May 13, 1939, for providing random noise distribution over the entire speech band.

The various control circuits of Fig. 2, which will be described presently, control the transmission of energy from either the relaxation oscillator 47 or the noise source 55 or both together into the circuit 56 leading eventually into the resonant circuits 57 and 58 which are provided for shaping the waves in a manner generally analogous to that in which the mouth cavity shapes the waves passing through the mouth in talking. The resulting shaped waves after transmission through these resonant circuits are impressed on the outgoing circuit, shown as comprising amplifier 59 and loud-speaker 60.

The resonant circuit 57 includes inductance coils 61, 62, 63 and 64 and condenser 65. Damping resistance may be included as indicated. The frequency of resonance of this combination is varied by varying the inductances. This is accomplished by varying the saturation of the coils on which these inductances are wound. Battery 66 is a biasing battery for determining the normal point on the permeability curve. Variable direct current is supplied in a manner to be described across the circuit terminals 67 and this variable current flowing through the inductance coils changes the degree of saturation and the inductance, thus varying the tuning of the resonant circuit. Condensers 98 and 99 are stopping condensers.

The resonant circuit 58 is constructed in the same manner and is arranged to be controlled simultaneously by the same control current. Resonant circuit 57 may be, for example, normally resonant to a low frequency and variable over the lower frequency portion of the speech band while circuit 58 may be normally adjusted to a relatively high frequency and variable over the upper portion of the frequency band as disclosed in my prior application Ser. No. 324,286, above referred to. By means of the resonant circuits 57 and 58 waves from the sources 47 and 55 are molded or shaped to reconstruct the speech which is supplied to the output 59, 60. There may be additional resonant circuits provided and controlled in the same manner and this is indicated in the drawings by the dotted conductors.

The operation of the circuit of Fig. 2 will now be described. When voiced waves are received from the analyzer, the variable direct current from the fundamental frequency channel passes through the filter 46 and controls the fundamental frequency of the relaxation oscillator 47 to accord with that of the impressed speech. These voiced waves result, in the manner above described, in the actuation of the relay 76 of Fig. 1 and in the connection of the source 78 to the line 14. Current of this frequency is selectively passed through the band-pass filter 80 of Fig. 2, rectified at 81, smoothed by the low-pass filter 82 and actuates relay 50 which attracts both of its sets of armatures. The waves from the oscillator 47, therefore, take the path through amplifier 53, front contacts of relay 50 and pass into circuit 56 leading to the resonant circuits

57 and 58. Relay 53 is closed whenever speech of any type, voiced or unvoiced, is received over the line and therefore the front contacts of this relay are now closed. (Relay 53 is operated from rectified current received through rectifier 84 and smoothing low-pass filter 85.) Current in the output of filter 46 passes in part into the circuit branch 86 and this has the effect of biasing the amplifier 54 beyond its cut-off point so that none of the output of the noise source 55 is allowed to be transmitted through amplifier 54. This biased voltage is impressed across the grid resistor 87.

The amplifier 95 performs no function under the conditions assumed that only voiced energy is being received. It may be stated, however, that the bias applied to resistor 88 of amplifier 95 from circuit branch 86 is opposite to that applied to resistor 87 and is such as to enable amplifier 95 to transmit waves from the noise source 55 into its output circuit. These waves are rectified at 89 and energize relay 51. This has no effect at this time, however, since relay 50 is assumed energized and is holding open the back contacts of its upper pair of armatures. It will be noted, also, that the output of amplifier 95 finds no closed path into circuit 56 through buffer amplifier 90 and contacts of relay 50, since these contacts are assumed open.

The frequency of the control generators in the analyzer of Fig. 1 will be assumed for illustration to be 105 cycles. The currents of these frequencies as received in the circuit of Fig. 2 are selectively passed through the high-pass filter 92, rectified at 93 and the rectified current is smoothed in filter 94 providing a direct current varying at no greater rate than 25 cycles per second, which current controls the resonance of the circuits 57, 58, in the manner that has been described. These circuits act similarly to the resonant chambers of the mouth and mold the waves to reproduce the voiced sounds assumed to be coming in over the line.

The next assumption will be that an unvoiced sound is coming in from the analyzer. This requires that the noise source 55 be connected to the circuit 56 but that the relaxation oscillator 47 be disconnected therefrom. The absence of a current in the output of filter 46 removes the blocking bias from the amplifier 54, allowing currents from source 55 to be transmitted directly to the circuit 56. Absence of any current in the output of filter 46 leaves the normally blocking bias applied across resistor 88 of amplifier 95, preventing passage of currents from noise source 55 through amplifier 95. Absence of any voiced speech results in relay 50 being deenergized. This, however, has no effect since relay 51 is also deenergized so that no circuits are closed through the contacts of the latter relay and no circuits controlled by contacts of relay 50 are closed by relay 51. Moreover, in the absence of current in the output of filter 46, the relaxation oscillator 47 is biased so highly negative that no oscillations are generated. The effect of the operations thus far described for the condition of unvoiced currents only coming into the analyzer is that the noise source 55 transmits some of its output into the circuit 56.

The unvoiced energy causes resonance indicating currents on the line which operate relay 53 as above explained, thus extending the circuit 56 to the shaping networks 57 and 58. The resonance of the circuits 57 and 58 is controlled by currents in the analyzer coming in through

filter 92, as previously described, and the currents resulting from the operation of the shaping circuits 57 and 58 are impressed on the output 60 and produce therein unvoiced sounds corresponding to those analyzed.

It will next be assumed that mixed sounds partaking of the nature of both vowels and consonants are coming in from the analyzer. These sounds, for example, are *v*, *z*, *zh* and *th* (then). Under these conditions, the relay 76 in the analyzer is deenergized and the fundamental frequency channel is energized as well as the appropriate relays in the analyzer circuit. Considering the effects in Fig. 2, the current in the output of filter 46 determines the fundamental frequency of the relaxation oscillator 47. Since there is no current of the frequency of source 78 from the analyzer, relay 50 is deenergized. The presence of current in the output of the filter 46 blocks amplifier 54 and unblocks amplifier 95. Relay 51 is, therefore, energized. Currents from both the relaxation oscillator 47 and the noise source 55 pass through the buffer amplifiers 90 and 91, respectively, through back contacts of relay 50 and into the circuit 56. Relay 53 is energized from line current and the control currents from the analyzer pass through the filter 92 and control the resonances of the shaping circuits 57 and 58.

When mixed sounds are produced in the case of human speech, the mouth opening is constricted tending to make the resonances less sharp. A corresponding effect can be produced in the electrical synthesis system by introducing electrical damping into the resonant control circuits 57 and 58. Such resistances are shown at 96 and 97, for example. Since the contacts of relays 51 and 59 in the circuit of these resistances are closed in the case of mixed sounds as described, these resistances are connected across the resonance control circuits 57 and 58 introducing damping into these circuits.

The explosives cause the operation of relay 24 in Fig. 1 to send current from source 22 over the line. This current is selectively received in Fig. 2 by band-pass filter 48, rectified at 68, smoothed at 69 and causes operation of relay 52 which moves its contact rapidly over the potentiometer resistance 52', thus producing abrupt changes in volume of the waves passing into the resonance circuits 57 and 58 to simulate the action of lips and tongue in making the explosive sounds.

The resonant circuits 57 and 58 and their mode of control, as shown, are in accordance with the preferred design. However, there is great latitude as to their construction and control and as to the number used. The showing is to be regarded as illustrative rather than limiting. In some cases simpler constructions may suffice. The circuits, as shown, permit almost any desired relation to be obtained between control current and valve of inductance, within wide limits.

Assuming that inductances 63 and 64 are on saturable cores and that inductances 61 and 62 are on saturable cores, saturated in the opposite direction, then if the control current increases the saturation in one pair of cores, it will decrease it in the other pair. This gives one inductance characteristic falling off with saturation and one increasing. By selecting these coils to start their saturation dropping at the desired point and to fall at the desired rate and have the proper amount of total inductance, it is possible in this way to get inductance characteristics with saturation that both drop off and later increase

as the saturating current increases. This is indicated in Fig. 3, where curve L' shows the manner in which the inductance of coils 61 and 62 varies with control current, and curve L'' is a similar curve for coils 63 and 64. The resultant is given by curve L'''.

By using more than two sets of inductances and by using them in parallel as well as series, any sort of inductance with saturation that is desired may be obtained within reason. The control current sets the first inductance at the value desired for the main resonance and the second inductance at a value to correspond to the resonant frequency paired with the main resonant frequency in the case of speech production.

In the circuit of Fig. 1, the recorder 100 is shown in position to be connected across the output of the analyzer when switch 101 is closed. This recorder may be used for monitoring or it may be used in lieu of the transmission line 14 for making a record which can be reproduced at some later time in a reproducer such as 102 indicated in Fig. 2. This reproducer may be connected into the synthesizer circuit by movement of switch 103 to close contacts leading to the reproducer. The output of the analyzer can be recorded on a record and reproduced from a record to very great advantage in view of the very limited frequency range required. On the basis of three channels operating at syllabic frequencies and using single side-band transmission, as indicated, the total frequency range would not be over 130 cycles or so at the most. Accordingly, the recorder could be driven at an extremely slow speed, say one-twentieth as fast as would be required to record a speech band of 3000-cycle band width. Wax records designed for a fifteen minute recording in the case of a 3,000 speech band would suffice to make a record twenty times that duration in the case of recorder 100, a duration corresponding to five hours. This would be valuable in the case of the storing of material for records. The recording may be done on wax or on film or on magnetic tape or in any other suitable fashion.

What is claimed is:

1. Means for analyzing speech to determine the frequency region of greatest energy per individual sound, a source of waves having frequencies distributed over the essential speech range, a modulating circuit comprising controllable resonance means, the frequency resonance of which is continuously variable, connected to receive the waves from said source, and means to control the resonant frequency of said resonance means in accordance with the speech frequency region of greatest energy as determined by said analyzing means.

2. In the production of artificial speech, means for analyzing input speech waves to determine the frequency band of principal resonance for each significant speech sound, means to generate waves covering the speech frequency range, a sound reproducer, means having controllable resonance, the resonant frequency of which may be continuously varied, connected between said generating means and sound reproducing means, and means controlled by said analyzing means for variably controlling the resonant frequency of said resonance means in accordance with said band of principal resonance as determined by said analyzing means.

3. In a system for artificial production of speech, sources of electrical waves having the characteristics of voiced and unvoiced sounds,

respectively, a sound reproducer energized from said sources, electrical resonance means connected between said sources and said reproducer for controlling the character of the reproduced sounds, said electrical resonance means being capable of continuous variation of resonance and simulating the effect of the resonant air cavities of the mouth, and means controlled by spoken sounds for controlling said resonance means.

4. In combination, an input for speech waves, a multiplicity of filtering circuits for subdividing said speech waves into frequency subbands, differentially operating means connected on the output sides of said filtering means for giving an indication of the subband containing maximum speech energy at a given instant, speech reconstructing means comprising electrical wave producing means and sound reproducing means actuated by the waves from said wave producing means, said reconstructing means including circuits of variable resonance for determining the frequency amplitude relations of the waves applied to said sound reproducing means, and means for controlling the resonance of said circuits of variable resonance from instant to instant in accordance with said indications.

5. In speech production, means to analyze input speech waves into voiced sounds, unvoiced sounds, mixed voiced and unvoiced sounds, and explosive consonant sounds, means to produce an indication of each of said analyzed sounds, including means to indicate the character of the voiced and unvoiced sounds, a source of electrical waves simulating voiced sounds, a source of electrical waves simulating unvoiced sounds, sound reproducing means energized from said electrical waves, circuits of variable resonance between said electrical wave sources and said reproducer, means variably controlling the resonance of said circuits in accordance with said voiced, unvoiced and mixed voiced and unvoiced sound indications, and means for abruptly changing the volume of reproduced sounds under control of said indications of said explosive consonant sounds.

6. The combination of claim 5 including means operating only in response to said indications of mixed voiced and unvoiced sounds for introducing damping resistance into at least one of said circuits of variable resonance.

7. In a speech analyzer, an input circuit for speech waves, a plurality of relatively narrow-band filters connected thereto, a pair of polarized relays each having one energizing winding connected to the output of one filter and a second energizing winding connected to the output of a filter passing an adjacent band, the two windings on each relay being mutually opposed in effect, whereby the relay energized is determined by the filter carrying maximum energy at a given instant, contacts controlled by said relays, and sound producer means controlled by said contacts.

8. In combination, a source of electrical waves of speech frequencies, a sound producer adapted to be energized therefrom, a resonance control system between said source and said producer, including inductive windings for determining the frequency of resonance, a source of input speech waves, and means controlled by said speech waves for varying the inductance of said windings.

9. The combination defined in claim 8 including means for producing a control current, means for varying a characteristic of said control current under control of said input speech waves,

and means for translating said variations in characteristic of said control current into variations in inductance of said inductive windings.

10. In artificial speech production, a source of electrical waves of speech frequencies, a sound producer energized thereby, means for varying the frequency amplitude relations of the reproduced sounds comprising a resonant system included in circuit with said source, said resonant system comprising a pair of inductances, 10

means for controlling the value of said inductances in accordance with a characteristic of a control current, means producing opposite, unequal variations in said inductances by the same variation in control current, whereby continuous variation in control current through a series of values of said characteristic in accordance with a given function is translated into a variation of inductance following a different desired function.

HOMER W. DUDLEY.