

May 27, 1941.

H. W. DUDLEY

2,243,089

SYSTEM FOR THE ARTIFICIAL PRODUCTION OF VOCAL OR OTHER SOUNDS

Filed May 13, 1939

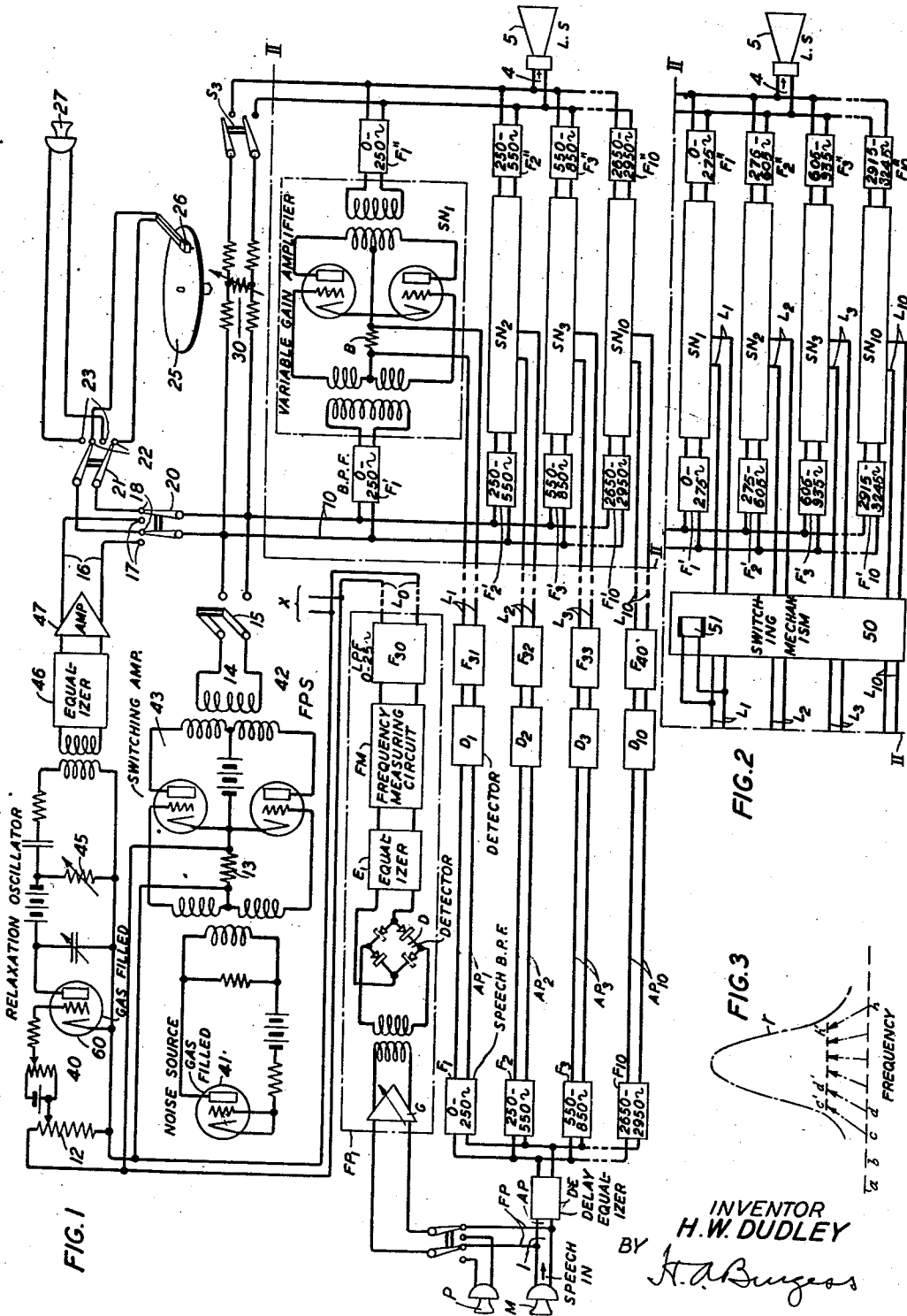
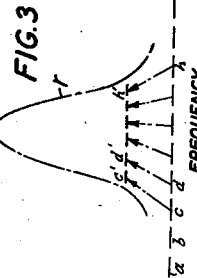


FIG. 2



INVENTOR
H. W. DUDLEY

BY *H. A. Burgess*

ATTORNEY

UNITED STATES PATENT OFFICE

2,243,089

SYSTEM FOR THE ARTIFICIAL PRODUCTION
OF VOCAL OR OTHER SOUNDSHomer W. Dudley, Garden City, N. Y., assignor to
Bell Telephone Laboratories, Incorporated, New
York, N. Y., a corporation of New York

Application May 13, 1939, Serial No. 273,429

26 Claims. (Cl. 179—1)

The present invention relates to wave transmission or control, involving waves of frequency band characteristics, for intelligence transmission, sound production or reproduction or other purposes.

The invention relates to the general type of wave control or wave production that is disclosed in my United States Patent 2,151,091, granted March 21, 1939, and for its general object it contemplates certain improvements or modifications of the basic system and method therein disclosed.

In accordance with the disclosure of my prior patent referred to, a wave having the characteristics of speech, for example, may be produced artificially using as raw material wave energy having a frequency distribution which may be uniform over a requisitely wide frequency range, and, therefore, lacking in character or intelligibility. The wave construction is accomplished by modifying such raw material wave energy in conformance with the fundamental pitch and energy distribution relations which characterize the speech or other wave to be produced. For best results the wave energy serving as the raw material comprises two types of waves, simulating respectively a buzzer-like tone and a hiss or unvoiced wave. An original speech or music wave may be analyzed for its fundamental characteristics of pitch and energy distribution. In a complete analyzer-synthesizer system, the result of the analysis is used to reconstruct the final wave out of the artificially supplied waves. For a more complete understanding of the systems and methods involved, reference is made to the patent disclosure.

While the system of my prior patent is susceptible of many types of uses and embodiments, one of its aims is to secure faithful transmission and reproduction of speech, music or other waves, although it is, of course, not limited to this aspect. The present invention may also be used similarly to the system of my prior patent, but the point of view of the present invention is directed more to securing a deliberate change of character between the analyzed waves and the reconstructed waves. Such change in character may be desirable for many purposes, such as rendering speech more understandable in the presence of noise or because of other conditions making reception difficult, or for purposes of imitation, impersonation, entertainment, research in speech or hearing, voice or musical training or instruction and kindred purposes.

The invention contemplates in its objects the realization of these and any other purposes for

which it may be desired to produce a controlled change in character between the analyzed wave and the reconstructed wave.

The invention contemplates specifically two general ways of changing the type or character of the reproduced wave. In accordance with one method, the wave used as raw material in reconstructing the output wave itself possesses or has imparted to it certain modifications or characteristics before it is acted upon by the products of analysis of the original wave or in such manner that it affects the character of the reconstructed wave. In accordance with the other mentioned method, a wave of either uniform energy distribution or non-uniform energy variation may be used as raw material, but the frequency regions chosen for reconstructing the output wave lack exact correspondence with those used in the analysis of the original wave and this results in a translation, change of voice or other effects to be described. These two methods of changing the character of the reproduced speech or music may be used separately or in conjunction with each other.

The nature and objects of the invention will be more fully apparent from the detailed description to follow, in connection with the accompanying drawing, in which:

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

Fig. 2 shows a modification which may replace the portion of Fig. 1 to the right of and below the broken line II—II; and

Fig. 3 shows a diagram of frequency relations to be discussed.

Referring now to the drawing, in Fig. 1, speech or music currents from line or circuit 1 energize a frequency pattern control circuit FP and an amplitude pattern control circuit AP. The frequency pattern control circuit, which comprises but one channel FP₁ discriminates as to the frequency pattern. This discrimination includes discrimination as to the fundamental frequency when there is one. The amplitude pattern control circuit branches into a number of channels, for example, ten channels AP₁ to AP₁₀ and determines what amplitude pattern we have. For simplicity, channels AP₄ to AP₉ are omitted from the drawing.

The information obtained from the speech analysis effected in these two circuits FP and AP is in the form of electrical currents which can be transmitted through any suitable transmitting medium, such for example, as lines L₀ to L₁₀, to the receiving or reproducing end of the

system. This transmitting medium may have a limited frequency range of transmission, of much less width than the frequency range of the speech signals to be communicated. If desired, it may be a single conductor pair or a radio link, the transmission through this medium then being, for example, on a carrier frequency basis in the general fashion shown in my above-mentioned patent.

At the receiving end of the system, in channel FP_1 the received waves act on an energy source of frequency patterns FPS so as to cause currents of the proper frequency pattern to flow from this source, and the received waves in channels AP_1 to AP_{10} are used to control shaping networks SN_1 to SN_{10} , respectively, to give the proper amplitude pattern to the power received by these networks from the energy source FPS . We then have our reproduction or reconstruction of the original speech signal for any further transmission in the ordinary manner, or for reproduction of sound by loud-speaker 5.

The system as shown uses a 275 cycle total transmission band in the transmission medium between the transmitting and receiving ends of the system, that is, in the lines L_0 to L_{10} . This 275 cycle band is on the basis of eleven channels, each of 25 cycle pass-band. Ten of these are for amplitude pattern control and the other one is for frequency pattern control. As brought out in my above-mentioned patent, such a system is adequate for good quality transmission of speech.

The frequency pattern control channel FP_1 will first be described with reference to its adjustment for use in natural reproduction of speech. It is a circuit for analyzing and reproducing the frequency spectrum of the source of energy in speech sounds. For speech applied to its input from line 1 it delivers an output wave that has discrete components and is of the same fundamental frequency as the input when a voiced speech sound is applied, and an output wave with a continuous spectrum when an unvoiced speech sound is applied. It performs three functions. First, at the transmitting or analyzing end of the system it derives from the speech signal the fundamental or vocal cord frequency and expresses this as a current, the amplitude or magnitude of which is proportional to the fundamental frequency. Next, at the receiving or reproducing end of the system it uses this current to control the frequency set up by a relaxation oscillator so as to get back a wave of the original fundamental frequency, rich in upper harmonics. Finally, it provides for another source of energy at the receiving end, having a continuous spectrum, when there is no fundamental in the speech. This condition occurs when sounds are unvoiced, as for example, in whispering and the unvoiced consonants.

From the standpoint of the source of acoustic energy the sounds of speech may be divided into three classes:

1. *Voiced sounds*.—For these sounds the acoustic energy is derived from the vibration of the vocal cords and the frequency spectrum is characterized by discrete components.
2. *Unvoiced sounds*.—For these sounds the acoustic energy is not derived from the vibration of the vocal cords but from the passage of an air stream through a restricted aperture or from the sudden stopping or starting of such a stream of air. The

frequency spectrum of such sounds is, in general, continuous and devoid of discrete components.

3. *Mixed sounds*.—For these sounds the acoustic energy is derived from both the above sources and the frequency spectrum consists of discrete components superposed on a continuous spectrum. The frequency pattern control channel FP_1 is a circuit that analyzes speech sounds on the basis of the above classification and which, within certain limits, delivers to a load a wave whose frequency spectrum is of the same class as that of the speech sound applied to the input of the circuit.

The circuit operates so that when no signal or an unvoiced speech sound (Class 2, above) is applied to the input, random noise from a gas-filled tube amplifier is supplied to the load. The spectrum of such random noise is a continuous one and so is similar to that of the source of acoustic energy in unvoiced sounds but any other source with a similar spectrum could be used. When a voice or mixed speech sound (Classes 1 and 3, above) is applied to the input, the noise voltage is removed from the load and a wave of the same fundamental frequency as the speech sound and any desired frequency spectrum is supplied to the load.

The analyzing circuit of the frequency pattern control channel FP_1 , or the portion of this channel at the sending or analyzing end of the system, will first be described. It comprises a detector D which may be, for example, a full-wave copper-oxide rectifier, an attenuation discrimination network or so-called equalizer E_1 having its loss increasing with frequency, a frequency measuring circuit FM and a 25 cycle low-pass filter F_{30} . The speech currents from line 1 are fed through the rectifier D , which feeds the equalizer E_1 , which in turn feeds the frequency measuring circuit FM . This frequency measuring circuit may be any suitable circuit for delivering through the low-pass filter F_{30} a direct current that depends on the number of reversals per second of direction of the voltage wave applied to this circuit and is independent of its amplitude as long as the amplitude exceeds a certain threshold value. For example, this circuit may have the form shown in Fig. 2, of Patent 2,183,248 of R. R. Riesz issued Dec. 12, 1939. This form is preferred on account of its being stable and free from singing, free from false operation at high frequencies, positive in action upon application of the input wave, and economical of plate battery power. In order to have the output current of the frequency measuring circuit controlled by the fundamental of the voice, the rectifier D modulates the various harmonics to give a strong fundamental and the harmonics are suppressed by the equalizer E_1 . The equalizer may be any suitable network having its loss increasing with frequency so as to insure that the fundamental frequency, which may vary for example from about 80 to some 300 cycles, comes out at a high power level compared to any upper harmonics that may be present. The equalizer may practically cut off transmission above a frequency in the neighborhood of 300 cycles, for example. For practical purposes the attenuation discrimination of the equalizer purifies the fundamental tone though it may vary substantially more than an octave. The level of the unvoiced sounds must be ad-

justed to a value too low to cause operation of the frequency measuring circuit. If desired, for this purpose a voice amplifier G, with its gain adjustable, may be provided, for example in the channel AP₁ as shown in Fig. 1. The direct current delivered by the frequency measuring circuit through the low-pass filter F₃₀ may be made substantially directly proportional to the fundamental frequency applied to the frequency pattern control channel from line 1.

The direct current delivered by the low-pass filter F₃₀ which may be substantially directly proportional to the fundamental frequency applied to the frequency pattern control channel from line 1 is transmitted through line L₀ to the energy source of frequency patterns FPS. The latter comprises a relaxation oscillator 40 and a noise source including gas-filled tube 41 followed by switching amplifier 43.

The relaxation oscillator comprises gas-filled tube 60 together with circuit elements as described more fully in the Riesz patent above referred to. This oscillator generates a wave which is rich in harmonics and the fundamental frequency of which is controlled by the voltage appearing across resistance 12 due to current flow in the line L₀.

The noise source 42 comprises gas-filled tube 41 having its grid tied to the cathode and having suitable resistances in its plate circuit as shown together with a plate battery. This tube while shown as a triode may advantageously be a multi-grid tube, such as a pentode or tetrode. It is found that this type of circuit produces a continuous energy spectrum of noise in the audio range. If desired, an equalizer or amplitude limiter (not shown) may be included between tube 41 and switching amplifier 43 to make the output flat over the frequency band. Instead of the type of noise source shown, a resistance source of noise may be used as disclosed in the Riesz patent referred to.

The function of the switching amplifier is to determine whether or not the continuous spectrum of noise is permitted to pass through to the circuit 14. This is accomplished under control of currents supplied across the grid resistor 13 from the control line L₀. The bias voltage on the switching amplifier in the absence of any voltage drop across resistance 13 is insufficient to block transmission. When voiced sounds are impressed on the system the voltage developed across resistance 13 is sufficient to block transmission.

Conversely, the initial grid bias on the relaxation oscillator 16 is sufficiently negative so that in the absence of voltage applied across resistance 12 the tube will not oscillate. When a voiced wave is applied to the frequency measuring circuit the resulting direct current voltage across resistance 12 is of the right magnitude and sign to cause the tube to oscillate at some low frequency for a weak input current and at higher frequencies for stronger input current.

The equalizer 46 renders all of the harmonic components of the current wave of the relaxation oscillator 40 equal in amplitude. This may be followed by amplifier 47, if desired.

When the switch 15 is closed and the switch 20 is in its left-hand position to make contact with terminals 17 both the noise source and the relaxation oscillator are connected to the synthesizing portion of the system through circuit 10. This is in the condition for natural reproduction of speech or music. In this condition switch S₃ (its

purpose will be described later on) is assumed to be open.

The amplitude pattern control channels are circuits which at the transmitting or analyzing end of the system measure how much power there is in the speech signal in chosen small frequency bands and transmit this information by control currents to the receiving or reproducing end where the output of the energy source of frequency patterns FPS appearing in circuit 10 is shaped accordingly. For transmitting a speech frequency range from 0 to 2950 cycles, for example, the speech bands chosen may be, for instance, one band from 0 to 250 cycles and nine adjacent bands each 300 cycles wide, starting at 250 cycles. These bands are selected by filters F₁ to F₁₀ in the amplitude pattern control channels AP₁ to AP₁₀, respectively. Thus, of these amplitude pattern control channels used to transmit information about the amplitude pattern, the channel AP₁ transmits information about the amplitudes in the speech range 0 to 250 cycles, the channel AP₂ transmits information about the amplitudes in the speech range 250 to 550 cycles, the channel AP₃ information about the amplitudes in the range 550 to 850 cycles, etc.

Considering channel AP₁, for example, the output from the 0 to 250 cycle speech band-pass filter F₁ is fed to detector D₁, which may be, for instance, like the detector D. The syllabic frequencies in the output from the detector are passed through a 25 cycle low-pass filter F₃₁ and the resulting variable direct current is passed through line L₁. This variable direct current is then applied to a biasing resistor B to give a grid bias to a signal shaping network or push-pull variable gain amplifier SN₁, which accordingly varies its gain in amplifying the waves received from the energy source of frequency patterns FPS through 0 to 250 cycle speech band-pass filter F₁', so that the average power in this band of waves varies in accordance with the average power in the corresponding band of the speech signals. The energy from the amplifier SN₁ is then fed through a 0 to 250 cycle speech band-pass filter F₁' to the speech receiving circuit 4, where it is combined with the outputs from nine other speech band-pass filters (of channels AP₂ to AP₁₀) to give a reproduction of the original speech signal.

It will be understood that in the description thus far the filters F₁' and F₁' may have the same pass-band as filter F₁, and that channels AP₂ to AP₁₀ may be like AP₁ except as to frequencies involved. However, a modified arrangement in which the pass ranges are not the same will be described later on.

In the circuit design it is desirable to have the delay in the frequency pattern control channel and in all of the amplitude pattern control channels the same. If the frequency pattern control channel FP₁ tends to have more inherent delay than the amplitude pattern control channels, it is desirable to introduce a certain amount of delay in the amplitude pattern control circuit AP as is indicated by delay equalizer DE in Fig. 1.

The output of the filters F₃₀ to F₄₀ is unintelligible, rendering it difficult for unauthorized persons to tap the circuit and affording a high degree of secrecy in either wire or radio transmission. The waves transmitted, through the transmitting medium, such as lines L₀ to L₁₀, directly contain no intelligibility whatever.

Instead of transmitting the currents from filters F₃₀, F₃₁ . . . F₄₀ by separate lines as shown,

they may be sent through any other type of channel such as carrier channels or on a time division basis, as in my prior Patents 2,151,091 and 2,098,956 November 16, 1937, or in other suitable manner.

In the operation of the system, the ten circuits $SN_1 \dots SN_{10}$ have their tubes biased so that when no voltage is developed across resistance B, the circuits are blocked. Thus even though the noise from circuit 14 is impressed on these circuits it does not get through them in the absence of applied speech currents.

When speech or other sounds are directed against the microphone M in circuit 1 the sound is analyzed for its fundamental frequency and for the flow of energy in each of the ten frequency ranges passed by filters $F_1 \dots F_{10}$ in the manner described. The currents which represent these energy flows control the transmission, through the $SN_1 \dots SN_{10}$ circuits, individually, of the waves from the generating system FPS. These waves, in turn, are of the type to give a voiced sound or of the type to give an unvoiced sound, as determined by the frequency pattern channel. The result is a reconstructed speech or sound corresponding to the original.

The present invention contemplates the use of other sources of waves for purposes of reconstruction of sounds. For example, the drawing shows a phonograph 24 with record 25 and pick-up 26 as one source of waves. This source is substituted for the relaxation oscillator 40 when switch 20 is thrown to its right-hand position to open contacts 17 and close contacts 18, and when switch 21 is in position to close contacts 22. At 27 is shown a microphone for picking up sound waves such as speech, music or other sounds, and may be substituted for the record by shifting switch 21 to its alternate position to close contacts 23. The switch 15 may be either closed or open depending on whether the noise waves are to be used or not. An adjusted amount of the waves from the record 25 or microphone 27 can be passed directly to the output circuit 4 by closing switch S_3 and adjusting pad 30.

The invention contemplates use of a variable gain amplifier like SN_1 in the leads from the microphone 27 and from the reproducer 26 and connected to channel L_0 at terminals X in the same way that SN_1 is connected to channel L_1 , so that sounds from the phonograph or microphone 27 are suppressed when only unvoiced sounds are spoken into microphone M. This variable gain amplifier is preferably adjusted to have an abrupt cut-off so that it is unblocked and opened for transmission in response to even a small amplitude current in line L_0 , the amplitude modulation occurring principally in the SN circuits of the synthesizer. This arrangement permits the noise to be applied through switch 15 by itself for the unvoiced sounds, the phonograph sounds being cut off momentarily just as the relaxation oscillator energy is suppressed when only unvoiced sounds are to be reproduced, when the system is used for talking. The control of the noise source from line L_0 is unaffected by this circuit modification. The relaxation oscillator is without effect since switch 20 is in its right-hand position.

Some very interesting and novel effects can be obtained by use of the substitute source 25 or 27. For instance, if an instrumental rendition of the air of a song is used in circuit 10, and if the words of the song are spoken or sung into the microphone M, the reconstructed sounds appear

as a song with instrumental accompaniment or as a singing instrument. With switch S_3 open, no sounds emerge except when sounds are applied at M. With switch S_3 closed, pad 30 may be adjusted to give a background simulating instrumental accompaniment. With switch 15 open all of the reconstructed sounds must come from source 25 or 27 and in the case of words spoken into M the articulation is less perfect than if switch 15 is closed. However, the provision of more filters in the AP lines to further subdivide and extend the range would increase the fidelity of the sounds at M without the use of the source 42.

In discussing the different effects obtainable, it will be convenient at this point to adopt terms which distinguish the waves applied to circuit 10 from those applied to circuit 1. It is seen that the waves applied to circuit 1 determine which of the SN circuits are unblocked and the extent to which they are individually unblocked. These waves will be referred to as the spectrum control waves or spectrum since they determine the amplitude pattern or amplitude distribution over the frequency band. The waves introduced at 10 will be called the pitch control or pitch waves since they furnish the actual frequencies passing through the SN circuits. For example, if a talker had a fundamental frequency of 220, the harmonics supplied from oscillator 40 would be multiples of 220 whereas another talker might have a fundamental, of, say 190 and the harmonics supplied from oscillator 40 would be multiples of 190. Or, record 25 or pick-up 27 in supplying music waves to circuit 10 would fix the pitch of the reconstructed speech. In either case the spectrum waves would not affect the pitch.

A few of the sound effects obtainable with this system will now be given by way of example, it being understood that these are but a few of the many possibilities.

The pitch may be supplied by one talker at 27 and the spectrum by a second talker at M. It is understood, of course, that both talkers may be from recordings, a suitable reproducer being used in place of microphone M.

The pitch may be supplied from a record, orchestra or other instrumental music and the spectrum from a talker at M. In this case the speech is automatically made into a song.

The pitch may be supplied from a singer, actual or recorded, and the spectrum from a second singer at M. This presents an interesting case since the singer at M may sing off pitch without affecting the result. Thus a poor singer at M can be made to appear as a good singer provided the good singer is supplying the pitch. Also, a talker or singer at M can supply a foreign accent, nasal twang or southern drawl to a good voice, such as that of a prima donna. This effect might be used to change a song sung in one language into a song sung in a different language.

The pitch may be supplied by one singer and the spectrum from a person whispering at M.

The effects of various animated objects may be obtained. For example, if any sound of sustained character, such as the rustling of leaves, Niagara Falls, the roar of surf, machinery noise, bird songs, airplane drone, rainfall, tap dancing, thunder, etc., is used as the pitch and speech is used as the spectrum, the result is articulated sound of the type used as the pitch control so that we would then have "talking leaves," "the voice of raindrops," etc.

Mixed music.—One voice or one instrument may be used for the pitch control and a second voice or second instrument for the spectrum with either the same or different tunes in the two cases.

Variation may be introduced into either or both of the pitch and spectrum waves to secure vibrato or tremolo effects.

In the modification in which a musical selection is sent from the record 25 or pick-up 27 into the synthesizer and words are sung or spoken into the microphone M, it might be thought that the pitch of the voice would need to correspond with the tone range of the music in order to produce the effect of a song. For example, if the music were predominately in the high register and the output waves are to be in the soprano range, it might be thought that a bass voice at M would be unable to control the proper SN circuits to modulate the music suitably. This is found, however, not to be the case. If a man's voice is analyzed, for example, a bass voice, it is found to have frequencies extending over as high a range as those of a soprano voice. For example, a bass voice contains harmonics of the order of 7,000 cycles or higher. The principal difference between a man's voice and a woman's voice is in the fundamental pitch. Probably the main reason why a woman's high voice sounds thinner than a man's low voice is that there are fewer harmonics in a given frequency range in the case of a woman's voice, since they are based on a fundamental frequency of about twice that of a man's voice. There is further the psychology of the lower frequency energy present in the man's voice. It is apparent, therefore, that a bass voice spoken into microphone M would contain the necessary frequencies to produce control currents in all of the lines $L_1 \dots L_{10}$ and to control the transmission of musical waves, in any part of the spectrum, applied to the circuit 10. Conversely, a woman's voice would also cover a sufficient range to actuate all of the synthesizing branches so that songs could be produced in any register.

When sounds from the record 25 or microphone 27 are used for reconstructing the waves, not only is the pitch obtained from these sources, but also enough of the spectrum is obtained to bring out its salient characteristics, such as the typical resonances of a violin or a horn. So long as this spectrum does not interfere too much with the spectrum patterns transmitted through the AP channels, intelligibility can be passed over the latter.

If it is desired to control only the pitch characteristics by means of a second sound input, the circuit may be used as first described for speech transmission, except that the FP channel is disconnected from microphone M and connected to a second microphone P into which the pitch control sounds are directed. This offers the possibility of producing a hybrid voice effect in which the second voice controls only the pitch of the first voice at M. Also a violin or other means may supply a tune to microphone P while a talker speaks words into M in proper cadence. The result is a song. Other effects may be obtained by varying the method.

In the circuit of Fig. 1 the filters appearing in any one channel, such as L_1 , are all indicated as having the same pass range, for example, 0 to 250 cycles per second. The invention contemplates, however, the use of filters in the synthesizer having different pass-bands from those in

the analyzer. This is indicated in Fig. 2 in which each of the filters $F_1' \dots F_{10}'$ and the corresponding filters $F_1'' \dots F_{10}''$ have pass ranges different from the corresponding filters $F_1 \dots F_{10}$. In the case of speech this has the effect of changing the character, for example, from that of a man's voice to that of a woman's voice or vice versa. For example, the pass ranges of the filters in the synthesizer may be moved upward in the frequency scale by 10 per cent as indicated in the following table in which the total utilized range is assumed to be greater than that shown in Fig. 1.

Analyzer bands	Synthesizer bands
0-250~	0-275~
250-550	275-605
550-850	605-935
850-1150	935-1265
1150-1450	1265-1595
1450-1750	1595-1925
1750-2050	1925-2255
2050-2350	2255-2585
2350-2650	2585-2915
2650-2950	2915-3245
2950-3500	3245-3850
3500-4500	3850-4950
4500-6000	4950-6600
6000-7500	6600-8250

In addition to changing the transmission ranges of the filters it will ordinarily be desirable to change the pitch of the voice, which can be done very simply by shifting the movable slider along the resistance 12 of the relaxation oscillator so as to vary the ratio between the generated frequency and the amplitude of the frequency control wave. A woman's voice has a fundamental frequency about twice as high as a man's voice. This means that there are twice as many harmonics in a man's voice of given frequency band as in a woman's voice of the same band. Changing the transmission bands of the filters alone would produce the effect of a woman's voice of low pitch, such as contralto, if the talker were a man. By changing the pitch as well as the pass-bands of the filters a simulation can be obtained. If, on the other hand, the synthesizer filters had their pass-bands shifted downward by, say, 10 per cent, and the pitch were lowered a suitable amount, the effect would be to translate a woman's voice into a man's voice.

To obtain even better simulation of a female voice, for example, measurements can be made for all sounds and the needed shifts in response frequencies provided for, including upward shifts at certain frequencies and downward shifts at others. These shifts are made merely by adjusting the pass ranges of the synthesizer filters. Such changes can be carried out to any desired detail. Similar changes can be made for translating from an adult voice to a child voice or from the human voice to speech having the characteristics of animals, either very small animals or large animals. For example, a parrot's speech is not only high pitched but confined to a relatively narrow frequency range because of the smallness of the talking apparatus which determines the resonance. If an attempt is made to convert from human speech to that of a parrot by merely tuning the talking circuit, it is found that the articulation is very poor and that the amplitude in the resonance region is unduly high.

Fig. 3 represents by the series of dashes a, b, c , etc., the band-pass ranges of the analyzer filters suitable for human speech. The dotted curve r indicates in a general way what the resonance

characteristics of a parrot or other small animal may be. By shifting the pass ranges of the synthesizer filters to the frequency positions indicated at *c'*, *d'*, etc., and making these bands narrower so that a greater number of them can be included within the resonance range of curve *r*, and by using suitably higher pitch, the output speech may be made to appear like that of a parrot or small animal and enough frequency components of the original speech will be utilized to permit of good articulation.

On the other hand, if human speech were to be made to simulate the speech of a wolf, lion or giant, or other large being where the resonances are broad and low, the pass ranges of the synthesizer filters would need to be shifted downward in the frequency spectrum and in some cases at least broadened. In some cases it might be necessary to shift some of the ranges downward and others upward to meet the requirements of a double humped resonance curve. Measurements may be made in any case to determine the type of shift necessary. As a further example, the following table indicates the kind of shifts that might be necessary to imitate the speech of a small animal.

Analyzer bands	Synthesizer bands
0-250~	0-500~
250-550	500-1000
550-850	1000-1500
850-1150	1500-1700
1150-1450	1700-1900
1450-1750	1900-2000
1750-2050	2000-2100
2050-2350	2100-2200
2350-2950	2200-2400
2950-3500	2400-2600
3500-4500	2600-3200
4500-6000	3200-4500
6000-7500	4500-6000

Various other effects may be obtained by controlling the spectrum with or without controlling the pitch. For example, a bass voice may be translated to soprano and vice versa. Thus a soprano singer of great musical ability would be enabled to sing in the bass register or some other register with the same characteristic quality.

In Fig. 2 the control channels are indicated as being carried through a switching mechanism designated merely by the box 50 in the drawing. This is to indicate that the various channels can be interchanged at will by the switching mechanism. For example, incoming channel 1 can be connected to outgoing channel 9 or any incoming channel can be connected to any outgoing channel by means of the switching mechanism. Some interesting effects are possible by this means. For example, if channels 1, 2 . . . 10 are active and the sound *e* as in *meet* is spoken into the microphone M it is heard from the loud-speaker as the same sound. With only channels 1 and 9 operative the sound remains practically the same. If, however, channel 9 is disabled and channel 1 from the analyzer connected to channel 2 in the synthesizer, then the sound *e* spoken into the microphone is transformed into the sound *o* as in *note* from the loud-speaker. With channel 1 going to channel 3 the sound becomes *o* as in *not* and with channels 1 and 9 going to 3 and 5, respectively, the sound becomes *aw* as in *haw*. It is actually possible, by manipulating the switching mechanism, including the shifting of energy sources, to

cause the word *yes* spoken into the microphone M to emerge from the loud-speaker as *no*, the word *we* to emerge as *you*, etc. This suggests the desirability of an automatic control of the switching mechanism to produce predetermined types of translation. The currents in the individual lines *L*₁, *L*₂, etc. are in the form of slowly varying direct currents suitable for controlling relays. This offers the possibility of operating relays in the switching mechanism to make predetermined changes in the circuit connections of the individual channels under control of currents in these channels. The relay 51 is shown controlled directly from the channel *L*₁ and is symbolic of any relay switching mechanism necessary for carrying out the desired switching functions. For example, relay 51 may be arranged to switch channel *L*₁ to one of the other channels leading to a higher frequency filter *F*₂, *F*₃, etc., so that when low frequency energy of sufficient magnitude appears in channel *L*₁, it causes the production of higher frequency components in the reconstructed speech. Various modifications may be made in the character of the switching and in the manner of controlling the relays governing the switching.

Where a microphone pick-up has been disclosed it will be obvious that any other type of device for setting up electrical waves having sound wave characteristics is contemplated, including record reproducers, photocells, induction, dynamic, capacitive pick-up, etc. Likewise, in place of reproducing the sounds as at 5, they may be recorded, broadcast, transmitted over lines or utilized otherwise. All types of records are contemplated including film, disc, etc.

The invention offers valuable possibilities in re-recording. For instance, if a record is played into microphone M and if the output energy in circuit 4 is used to make another record, the translation from the first to the second record can be modified in a wide variety of ways, by use of the pitch microphone P, the microphone 27, or the record 25. The nature of the effects produced has been discussed above. In talking movies, these possibilities may sometimes save retakes, or may be used variously, as in a series of steps, to work up the desired type of final sound record.

The shifting of the pass ranges of the reconstructing filters relative to the analyzer filters may be used along with the alternative sources of pitch control waves 25, 27, etc. to secure desired effects.

It is recognized that the system disclosed herein is susceptible of wide variation and modification of an obvious nature in view of the examples given. All such variations and modifications are intended to be protected by this application. Moreover, the various examples, circuit details, frequency ranges and sound effects disclosed are to be construed as illustrative rather than limiting and the scope of the invention is defined in the claims which follow.

What is claimed is:

1. In combination, a source of music-representing waves, a source of speech waves, means to derive from each of a plurality of frequency subdivisions of the speech waves an index of the energy content of such frequency subdivision varying with time, means to modulate said music-representing waves in a plurality of circuits in accordance with said indices, and means to translate the resulting modulated wave energy into sound energy.

2. The method comprising subdividing a sound-representing wave into a number of relatively narrow frequency bands, subdividing a second wave of band characteristics into a corresponding number of frequency bands certain of which are of different band width from the corresponding bands of the first wave, deriving an index of the energy content of each of the subdivided bands of the first wave, and controlling transmission of each of the subdivided bands of the second-mentioned wave in accordance with a different one of said indices.

3. The method of producing vocal modulation of a musical wave comprising subdividing into narrow frequency bands the energy spectra of both the vocal wave and the musical wave, deriving an index of the energy content, varying with time, of each of the subdivided vocal wave bands, reproducing sound from the subdivided bands of the musical waves, and controlling the sound reproduction from each such subdivided band of the musical waves in accordance with a respective index so derived.

4. The method of reproducing speech with changed voice comprising analyzing the speech waves to determine their energy content in each of a plurality of distinct frequency regions, generating waves having a frequency band extending over the range of the sounds to be produced, and varying the energy content of each of a plurality of distinct frequency regions of said generated waves in accordance with the energy content of a respective region of the analyzed waves, the distinct frequency regions of the generated waves so controlled being displaced in frequency with respect to the respective distinct frequency regions of the analyzed waves.

5. The method of reproducing speech with changed voice comprising analyzing the speech waves to determine their energy content in each of a plurality of distinct frequency regions, generating waves having a frequency band extending over the range of the sounds to be produced, and varying the energy content of each of a plurality of distinct frequency regions of said generated waves in accordance with the energy content of a respective region of the analyzed waves, the distinct frequency regions of the generated waves so controlled being of respectively different frequency width from the corresponding distinct frequency regions of the analyzed speech.

6. In combination, a source of waves of a range of acoustic frequencies, a second source of waves of a range of acoustic frequencies, a set of filters for subdividing the waves from the first source into relatively narrow frequency bands, a set of filters for subdividing the waves from the second source into a corresponding number of relatively narrow frequency bands certain of which differ in width from the corresponding bands into which the waves from the first source are divided, means determining the energy content of each of the subdivided bands from said first source, means arranged to respond to the waves of the subdivided bands from said second source, and means for controlling the response of said second means to each of said subdivided bands from said second source in accordance with the energy content of the corresponding subdivided band from said first source.

7. In combination, means for analyzing speech waves to determine their fundamental frequency and their energy rate in each of a number of frequency regions, means for generating waves extending over a wide range of acoustic frequen-

cies comprised of a fundamental frequency and harmonics thereof, means causing the fundamental frequency of the generated waves to be different from the fundamental frequency of said speech waves but definitely related thereto, means producing sounds from said generated waves, and means controlled by said several energy rates for determining the effectiveness of said means in producing sound from each of several frequency regions of said generated waves, different in frequency from the regions of the speech waves used to determine said energy rates.

8. In combination, means for producing the effect of a small animal talking under control of human speech, comprising means for generating acoustic frequency waves extending over the relatively narrow frequency range appropriate to said animal, means for subdividing said waves into a number of narrow bands, means for subdividing the human speech waves into a corresponding number of narrow bands, means determining the energy rate of each of said latter bands, and a plurality of means controlled individually in accordance with the energy rate of said individual bands for producing sound in accordance with each of said bands of said generated waves.

9. In combination, means for producing the effect of a giant or large animal talking under control of human speech comprising means for generating audio frequency waves extending over the relatively great frequency range appropriate to said giant or large animal, means for subdividing said generated waves into a number of relatively narrow bands, means for subdividing the human speech waves into a corresponding number of narrow bands, means determining the energy pattern of each of said latter bands, and a plurality of means controlled individually in accordance with the energy pattern of said individual bands for producing sound in accordance with each of said bands of generated waves.

10. The combination according to claim 8 including means to derive the fundamental frequency of said speech waves, and means to control the frequency of the generated waves thereby.

11. The combination according to claim 9 including means to derive the fundamental frequency of said speech waves, and means to control the frequency of the generated waves thereby.

12. A system comprising a plurality of wave transmission channels for respectively transmitting subbands of a frequency band of waves that represents a voice wave, means for producing a complex tone having its fundamental frequency in controlled relation to that of said voice wave but having the relative amplitudes of its components independent of the relative amplitudes of the components of said voice wave, analyzing circuits responsive to waves from said channels for determining the energy flow, varying in time, in each of said subbands, means controlled by said analyzing circuits for producing from the energy of said complex tone a similitude of said energy, varying in time, in said subbands, and means for interchanging the control between individual analyzing circuits and the means for producing said similitude of said energy in said subbands.

13. A system according to claim 12 including means controlled by said voice wave for actuating said means for interchanging said control.

14. In combination, a source of wave energy of band characteristic, a plurality of transmission control means operative at different frequency regions within said band, means actuating one of said control means in accordance with voice energy in a low frequency part of the voice range to control transmission of low frequency waves from said source, means actuating another of said control means in accordance with an intermediate range of voice energy to control transmission of waves of intermediate frequency from said source, means actuating another of said control means in accordance with a high frequency range of voice energy to control transmission of waves of high frequency from said source, means making certain of the frequency ranges whose transmission is controlled by said several transmission control means different in width from the corresponding ranges of said voice wave, and means for combining the effects of the transmitted ranges.

15. The method of simulating the effect of a talking object which gives off a characteristic sound of its own, such as a waterfall, machine, etc., which comprises subdividing said characteristic sound into a number of relatively narrow frequency subbands, subdividing speech waves into a corresponding number of relatively narrow frequency subbands, utilizing the energy in each of said subbands of the speech waves to control transmission of the corresponding subbands of said characteristic sound, and making certain of the frequency subbands of said characteristic wave of respectively different width from the corresponding subbands of said speech wave.

16. The method of simulating the effect of a talking image or cartoon comprising supplying a frequency band of waves appropriate in pitch and band width to said image or cartoon, subdividing said band of waves into a number of relatively narrow frequency subbands, subdividing speech waves into a corresponding number of relatively narrow frequency subbands, controlling the amplitude of each of said frequency subbands of said first wave in accordance with the energy in each of said subbands of said voice wave and combining the effects of the energies of the subbands of said voice wave after such control.

17. The method of producing the effect of a singing musical instrument comprising taking the characteristic wave given off by said instrument in the form of a musical selection, subdividing said wave into relatively narrow frequency subbands, subdividing a voice wave into a corresponding number of relatively narrow frequency subbands, determining the energy rate in each of said frequency subbands in said voice wave, controlling the amplitude of the waves in each of the subbands of said musical wave in accordance with the energy rate of the corresponding subband of said voice wave and combining the effects of the subbands of said musical wave whose amplitudes are so controlled.

18. The method of producing sound effects comprising taking the voice of a singer in the form of a musical selection, subdividing said wave into relatively narrow frequency subbands, subdividing another voice wave into a corresponding number of relatively narrow frequency subbands, determining the energy flow in each of said frequency subbands of said other voice wave, controlling the amplitude of the waves in each of the subbands of said musical wave in accordance with the energy flow of the correspond-

ing subband of said other voice wave and combining the effects of the subbands of said musical wave whose amplitudes are so controlled.

19. The method of enabling a speaker to change his voice comprising analyzing his voice wave to determine its fundamental frequency and the energy rate in each of a number of relatively narrow frequency subbands, generating waves covering the speech range under control of said fundamental frequency, subdividing said generated waves into a corresponding number of relatively narrow frequency subbands certain of which differ in band width from the corresponding frequency subbands of said voice wave, controlling the amplitudes of the waves in the various subbands of said generated waves in accordance with the energy rate of the corresponding subband of said voice wave and combining the effects of the waves of said subbands whose amplitudes are so controlled.

20. The method according to claim 19 comprising generating said band of waves from a fundamental frequency dependent on the fundamental frequency of said voice wave but different therefrom.

21. The method according to claim 19 comprising making certain of the subbands of the generated waves of different band width from the corresponding subbands of said voice wave and shifted in frequency with respect thereto.

22. Means for producing the effect of a singing musical instrument comprising means for producing waves characteristic of such instrument and representing a musical selection, an input circuit, an output circuit, a plurality of transmission control devices connected in parallel paths between said input circuit and said output circuit, means for subdividing said characteristic wave on a frequency basis into a corresponding plurality of frequency subbands, means to apply the waves in each of said subbands to a respective one of said control devices for transmission therethrough, means causing waves to flow in said input circuit representing words spoken in cadence with said musical selection, means subdividing said latter waves into a plurality of frequency subbands, means to determine the energy characteristic of each of said latter subbands, and means to control the transmission of the first-mentioned waves through each of said transmission control devices in accordance with the energy characteristic of a corresponding one of said subbands representing said spoken words.

23. In a sound producing system, means to produce waves of the frequency range of sounds to be produced, based upon a fundamental frequency, means comprising a speech input circuit for speech waves from one wave source to control the amplitude relations throughout said band in accordance with the characteristic energy rates of speech at corresponding frequencies of the speech band, and means comprising a second speech input circuit for speech waves from another wave source to control said fundamental frequency independently of said first speech wave source.

24. In a sound system, two input circuits for sound-representing waves, a common output circuit, a source of complex tone waves based on a fundamental frequency, means for transmitting waves from said source into said output circuit, a plurality of amplitude pattern control channels for exercising control of the waves so transmitted in each of a number of frequency regions in accordance with frequency regions of the

sound-representing waves in one of said input circuits and a pitch control circuit for controlling said fundamental frequency in accordance with other sound-representing waves in said other input circuit.

25. A system comprising a plurality of channels for transmitting energy having syllabic time rates of change in the channels, respectively, a source of waves having a sound spectrum the relative amplitudes of the components of which are independent of the relative amplitudes of the components of said transmitted energy, means controlled by the energy in said channels for varying at said rates, respectively, the energy in different frequency regions of waves from said source, and switching means for interchanging said channels in their control of the energy in said different frequency regions.

26. The method of re-recording comprising reproducing sound-representing waves from a first record, subdividing said waves on a frequency basis into narrow bands, generating waves covering a sound range of frequencies, producing a second record from said generated waves, controlling the amplitudes of the components of the generated waves in different frequency regions in accordance with the energy flow in respective bands of said subdivided waves, and controlling the pitch of the generated waves independently of said reproduced waves from said first record in order to effect controlled changes in the re-recording.

HOMER W. DUDLEY.