

June 21, 1938.

H. W. DUDLEY

2,121,142

SYSTEM FOR THE ARTIFICIAL PRODUCTION OF VOCAL OR OTHER SOUNDS

Filed April 7, 1937

3 Sheets-Sheet 1

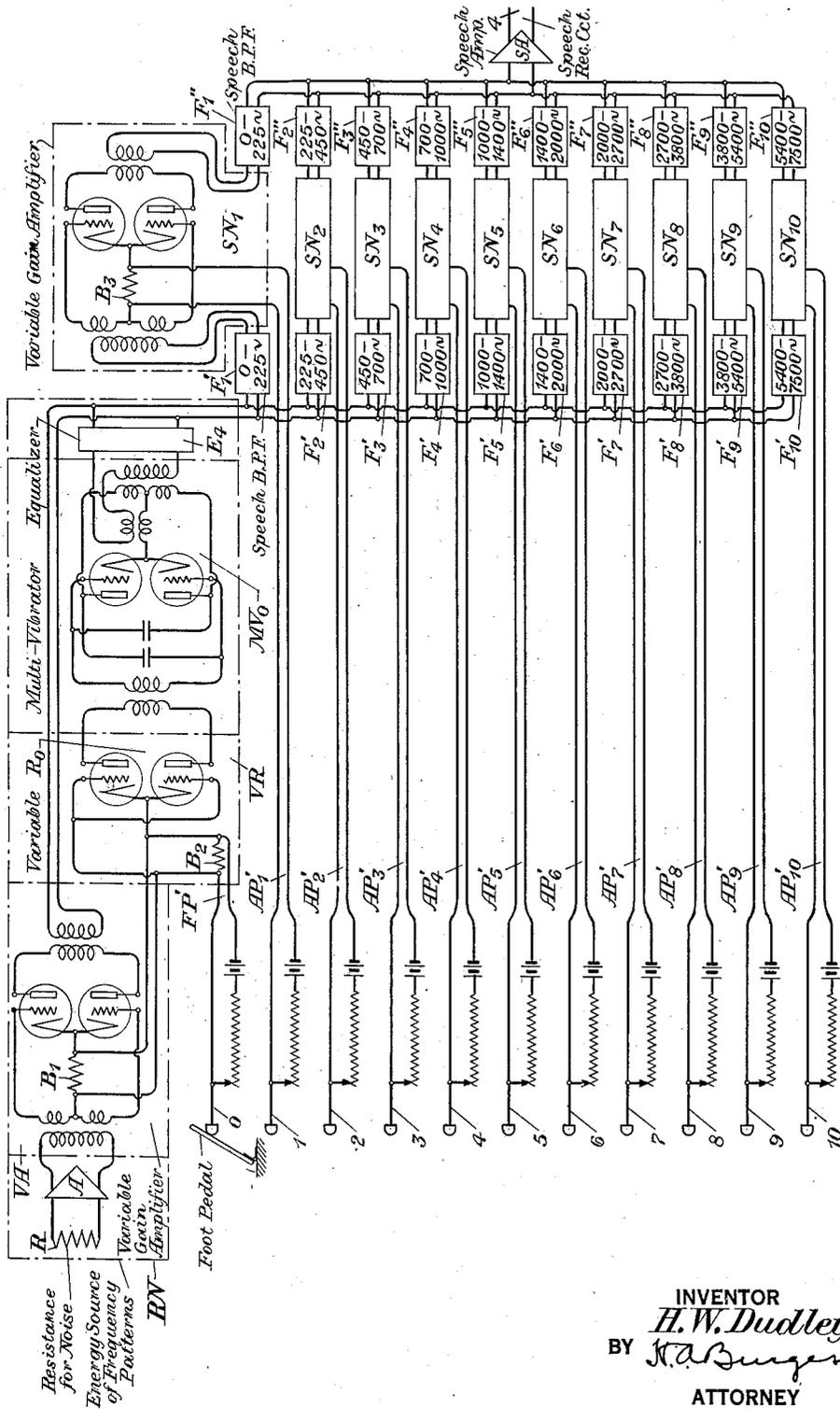


Fig. 1

INVENTOR  
*H. W. Dudley*  
 BY *J. A. Surges*  
 ATTORNEY

June 21, 1938.

H. W. DUDLEY

2,121,142

SYSTEM FOR THE ARTIFICIAL PRODUCTION OF VOCAL OR OTHER SOUNDS

Filed April 7, 1937

3 Sheets-Sheet 2

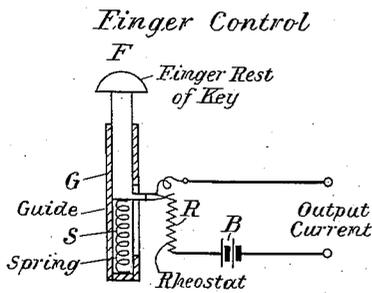


Fig. 2

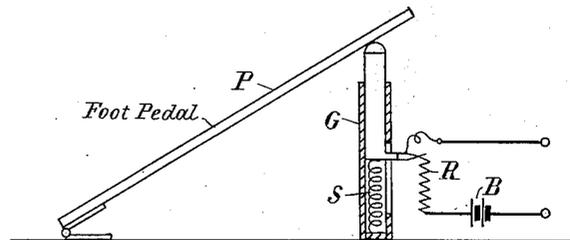


Fig. 4

Fingering Layout of Keys

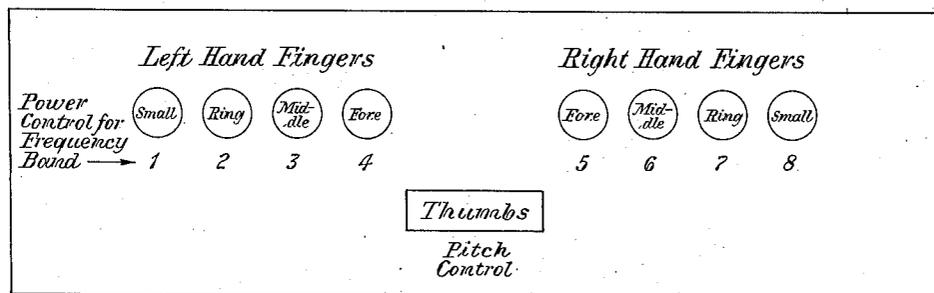


Fig. 3

Fingering Layout of Keys

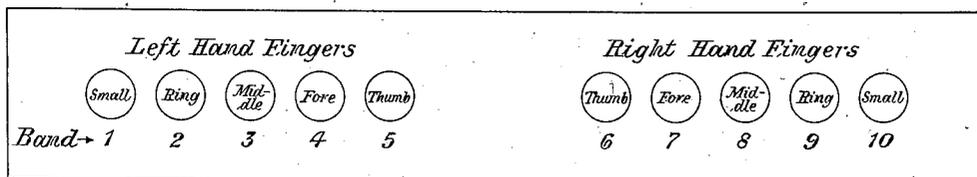
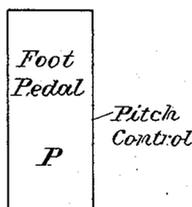


Fig. 5



INVENTOR  
*H. W. Dudley*  
 BY *A. A. Burgess*  
 ATTORNEY

June 21, 1938.

H. W. DUDLEY

2,121,142

SYSTEM FOR THE ARTIFICIAL PRODUCTION OF VOCAL OR OTHER SOUNDS

Filed April 7, 1937

3 Sheets-Sheet 3

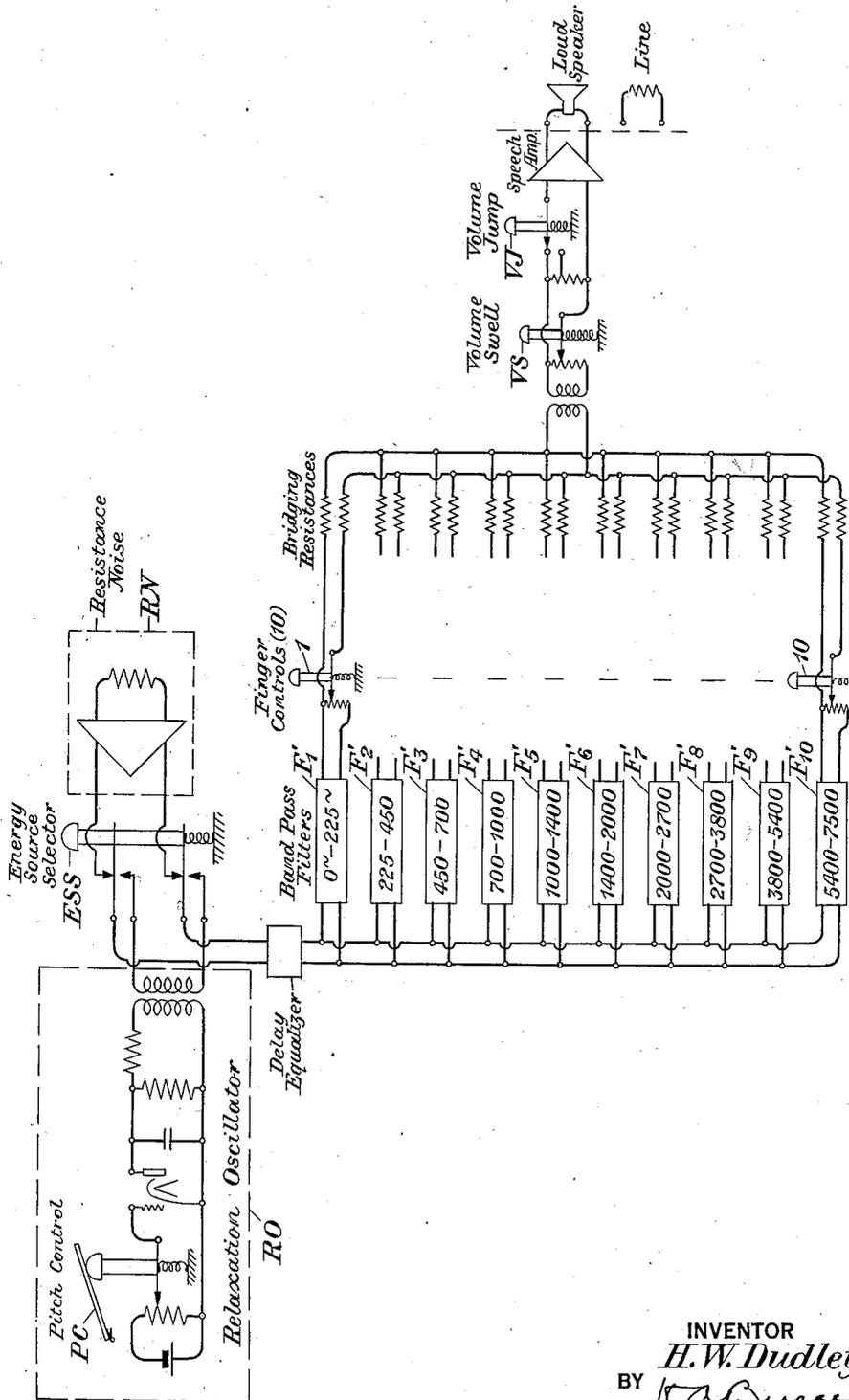


Fig. 6

INVENTOR  
*H. W. Dudley*  
BY *H. A. Burgess*  
ATTORNEY

# UNITED STATES PATENT OFFICE

2,121,142

## SYSTEM FOR THE ARTIFICIAL PRODUCTION OF VOCAL OR OTHER SOUNDS

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

Application April 7, 1937, Serial No. 135,416

20 Claims. (Cl. 179-1)

This invention relates to the artificial production of vocal or other types of sounds.

One of the objects of the present invention is to provide an arrangement for the synthetic production of speech or similar sounds and, particularly, to provide an arrangement of this type in which the desired vocal or other sounds may be produced by manual operation quite independent of any vocal control by the normal vocal mechanism of the human body.

In my prior application, Serial No. 47,393, filed October 30, 1935, use was made of the principle that speech may be resolved into invariable factors such as the vibrations of the vocal cords, and into variable factors such as the changes of pitch of the vocal cords and the various modulations effected by the lips, tongue, palate, etc. By the arrangement disclosed in such application, speech is instantaneously analyzed to determine the set of variable parameters which will define the unknown or variable elements of the speech signal. The fixed factors, such as the relatively high frequency vibration due to the vocal cords or to the hissing sounds of the air rushing through passages, are not transmitted. On the other hand, the variable parameters (which may correspond to the volume of energy in different frequency ranges of the voice, and in the variations of pitch of the voice sounds) are transmitted in separate channels to the speech synthesizer. The information received in the several channels is combined by the synthesizer with waves from local sources corresponding to the invariable characteristics of speech, to reproduce the original sound.

In accordance with the present invention it is proposed to use a portion of the synthesizer of my application, above referred to, in combination with certain manually operable equipment, to produce speech or other sounds artificially by purely manual operation independent of any of the parts of the human body which are normally used in the production of vocal sound. As an example, the invention may be embodied in an arrangement in which finger operated mechanisms are used to produce currents in the control channels of the synthesizer which correspond to those which would normally be received from the speech analyzer of my application, above referred to. If desired, of course, other parts of the body, such as the feet, may be used in effecting some of the controls. For example, the controls which determine whether the local oscillation generators shall produce a continuous frequency spectrum as in the case of a hissing or

unvoiced sound, or a discrete frequency spectrum as in the case of a voiced sound, may be operated by the feet. So, also, the control which determines the fundamental frequency in the case of a discrete spectrum, may be a foot control. However, other parts of the body are equally available for this purpose. For example, the controls just mentioned might be operated by pressure between the teeth of the upper and lower jaws.

Analyzing the vocal system of a man from the broad viewpoint of producing speech sounds, it is seen to be made up of two types of parts, (1) fixed, and (2) variable. The variable parts are considered here to be those that vary in position from sound to sound. Examples are the lips and teeth opening and closing, the tongue shifting forward and backward, the vocal cords varying in tension, and the uvula opening and closing the nasal passage. The term "fixed" is here used in its broadest sense. It not only includes parts that are not moved from sound to sound in speech, such as the nasal passages, pharynx and much of the larynx, but it also includes any fixity of feature. As an example, the fact that the vocal cords are always used in the voiced sounds is a fixed feature, as is also the fact that they always vibrate in the same buzzer-like way as regards the presence of a fundamental frequency and all of its overtones up to a large number greater than 30; the variation of the fundamental frequency, or pitch, of the vocal cords is, on the other hand, a variable feature, as stated previously.

The whole vocal system may be likened to a mechanical-acoustical oscillator with certain fixed circuits and certain variable mechanical elements. The part of the vocal system corresponding to the fixed circuits of the oscillator is the same from man to man. It is the same from sound to sound in the same man with the different elements taking on different values to produce the different sounds. To make the analogy just referred to more specific, consider that the vocal system is, in principle, like the ordinary electrical oscillator mounted in a box as a fixed piece of apparatus, the variability being obtained by switches for starting the oscillator and for choosing the desired inductances, by continuously variable dials for selecting the capacitance, and by step variable dials for adjusting the resistances controlling the output. With such an arrangement other features such as feedback may also be controlled.

Just as the oscillator when oscillating is still

essentially a fixed piece of apparatus to which variable controls of frequency, output and feedback, are applied, just so the fixed vocal system includes the condition of the vocal cords vibrating at an average or other specified steady rate, to which condition the controls for varying or modulating the generated vibrating signal in the production of speech can be applied. The importance of including a normal vibration of the vocal cords as a fixed feature is due to the fact that they vibrate on the average at a fundamental frequency of 100 to 150 cycles per second for a man and about twice this for women, whereas the variable controls specified can change only at rates ten or more times smaller than this in the case of men, and 20 to 40 times smaller in the case of women. Strictly, the fixed features include what may be described as a multi-oscillator source of energy rather than a single one, for not only are there the periodic oscillations produced at the vocal cords, but there are also non-periodic or random oscillations produced by the passage of air through restricted openings such as between the lip and teeth for the "f" sound, between the tongue and hard palate for the "sh" sound, between the vocal cords themselves for whispering, etc.

This differentiation between fixed and variable features also characterizes the type of modulated speech signal produced. In this case the fixed features correspond to the sustained oscillatory sound produced with the various elements or parts of the vocal system in an average or normal position. This means an average lip position, an average vocal cord tension, etc. The variable features correspond to the changing or modulating of the sound by varying the different elements from their average positions. It will be clear, therefore, that the fixed features appearing in a speech signal are oscillatory in nature and the variable features are modulatory.

An analysis of an oscillogram of a speech wave shows that there are variations from maximum swing in one direction to maximum swing in the opposite direction in .001 second or less. Yet these are oscillatory swings, for in the next period about .010 second later the same swing will still be found as an almost identical copy. If the slight change that occurs from period to period is followed up until it becomes great enough so that the original wave form is lost, this condition will be found to occur many periods later, oftentimes twenty or more, requiring a time of the order of .200 second. This latter type of change is the modulatory type of change. It is easiest seen in a single sound as the building-up of the peak amplitudes to a maximum and then a falling-off to zero again. In this case a complete change corresponds to a large part or all of a whole syllable and such change is therefore known as a syllable or syllabic frequency change.

From the foregoing it is evident that speech has a dual characteristic. On the one hand we have fixed parts or elements setting up oscillatory waves containing relatively high frequency patterns. On the other hand we have varying parts or elements setting up modulatory waves of low syllabic frequency pattern. An ideal arrangement for producing speech artificially would involve producing all of the fixed features by some artificial means and then applying thereto manually produced modulatory effects corresponding to the instantaneous positions of the

variable parts utilized in normal speech production.

It is well known that one set of parameters can be substituted for another without any loss of definition so long as the number of independent parameters remains unchanged. Any change from the simple ideal above mentioned generally leads to a large number of required parameters because the newly selected ones are not independent. However, this is not of much practical importance for, as will be pointed out later, a new set of parameters may be chosen which are quite simple and yet are almost as independent of each other as the fundamental parameters used in the normal production of speech.

As pointed out in detail hereinafter, the number of independent variables involved in the production of speech is small. That is, the number of movable or variable elements of the vocal system that are controlled as parameters to give the desired speech production, and are movable or variable substantially independently of one another by the muscles of the vocal system, is small. In other words, the number of variables or parameters that can be controlled substantially independently in speech production is small, being of the order of ten. Moreover, as indicated above and discussed hereinafter, for each of the physical elements the minimum time in which it can go through a complete cycle of change in position is not less than one-tenth of a second. Consequently, each independent variable has a fundamental frequency of not over ten cycles per second, while engaged in speech production.

As explained before, these speech defining signals may be any signals derived from speech signals, provided the derived signals give as many independent variable quantities or parameters as the number of independent variables involved in the production of speech. Furthermore, the chosen parameters need not be entirely independent, provided their number be increased sufficiently to make up for their lack of independence. For example, if the original speech band be divided into a sufficient number of sub-bands the chosen parameters may be merely the average amounts of power in the several sub-bands, as brought out in detail hereinafter.

There exists, then, in the actual production of a complex wave by the vocal system, a simple set of slowly varying elements or parameters (the independent variable elements referred to above of the vocal system), that determine the variable characteristics of the signal which are referred to above. And, to transmit information that will suffice for defining or reproducing the variable characteristics, it is unnecessary to transmit the fixed characteristics of speech, it being sufficient to transmit information defining the variations of any simple set of parameters derived from the complex speech wave and corresponding to the independently variable elements of the vocal system as regards number and independence, or as regards the number of degrees of freedom of variation.

In the arrangement disclosed in my earlier application, Serial No. 47,393, filed October 30, 1935, the foregoing principles are utilized in the transmission of a complex signal, such as speech, by sending to a receiving synthesizer variant information regarding the variable or unpredictable characteristics of the signal to be transmitted, instead of sending the complex signal wave itself. The waves thus transmitted can define the

signal precisely, as regards its unknown or variable characteristics, yet have small frequency range relative to the signal wave and at the same time be as short in duration as the signal.

5 In accordance with another feature of the invention of my prior application, above referred to, as applied to the transmission of speech, for example, the fact that much of the information ordinarily transmitted is of an invariable or  
10 predictable character, due to the general uniformity of the speech producing organs from person to person, is taken advantage of by reproducing such predictable information artificially at the receiving end of the transmission system,  
15 in order that it need not be transmitted from the sending end. Thus, effective use is made of the information or foreknowledge of the fixed or invariable characteristics of the signal source, with the result that the frequency band width of transmission can be reduced.

In one specific aspect, the invention of my prior application, above referred to, is a system in which a speech signal is analyzed for its fundamental frequency, and for the average power in  
25 properly chosen sub-bands of frequency, this information being transmitted and then used at the receiving end by means of a synthesizer, to fashion waves from a local multi-frequency source into a simulation of the signal. To fashion  
30 the simulation of the signal from the waves supplied from the local source, frequency sub-bands of these locally derived waves are selected which are, respectively, coextensive with the chosen sub-bands of the speech signal, and the average power  
35 in each sub-band of the locally supplied waves is varied in accordance with the power in the corresponding chosen sub-band of the signal wave. This variation is effected in response to the information transmitted from the sending  
40 end of the system regarding the average power in the chosen sub-bands of the signal wave.

Two types of frequency spectrum are used alternately in speech, (1) a continuous spectrum in the case of hissing or unvoiced sound, and (2) in  
45 the case of voiced sounds a discrete spectrum with a variable fundamental and with upper harmonics always present to a relatively high frequency. Hence, the local source provided at the synthesizer of my prior application, above referred to, preferably is such that the waves supplied by the local  
50 source can have either type of spectrum. The type is determined in response to the information transmitted from the sending end of the system with regard to the presence or absence of a fundamental frequency in the speech wave and the magnitude of any such fundamental frequency. In other words, if the fundamental frequency is present the discrete spectrum is generated by the local source, and if no fundamental frequency is  
55 present a continuous spectrum is generated.

Significant changes in the fundamental frequency of the speech sounds and in frequency distribution of power in speech, can take place only at a rate which is limited by the sluggishness of  
65 the muscles of the vocal system to less than about ten cycles per second (a frequency much lower than the fundamental oscillatory frequencies of vocal cords which range from about sixty cycles to in the neighborhood of five hundred cycles). It therefore results that the equipment required at the sending end of the system of my application, Serial No. 47,393, above referred to, for analyzing the speech signal as to its fundamental frequency, and likewise the equipment provided at  
70 the receiving end of the system for responding to

the transmitted indications as to the fundamental frequency of the speech sounds, need only be responsive on the line side to frequencies up to perhaps one to three times the frequency of ten cycles per second, just mentioned, depending on the accuracy desired in the transmission of the  
5 indications.

Moreover, the number of sub-bands analyzed for power content need not exceed five or ten, for example, to obtain high intelligibility; because,  
10 as indicated above and pointed out in detail hereinafter, the number of independent variables or parameters in speech is small, and the power in each sub-band is largely independent of that in the others, particularly as the distance between  
15 the mid-frequency bands is increased.

Such a system, then, analyzes the signal as to its fixed features and variable features. The fixed features include, (a) the existence of definite frequency sub-bands in which the power distribution is sensibly uniform; (b) the existence of a frequency spectrum that alternates from the continuous type of spectrum to a discrete type with varying fundamental and with all upper harmonics always present; and (c) the fact that time  
25 variations of the fundamental frequency and of the power in the frequency sub-bands occur only at syllabic frequency rates. The variable features include, (A) the magnitude of the average power in each sub-band, and (B) the nature of the signal spectrum (as to whether it is continuous or discrete and, in the latter case, as to what is the magnitude of the fundamental frequency).

Since there is foreknowledge at the receiving end as to the fixed features or characteristics of the signal, they can be supplied locally at the receiving end and it is unnecessary to transmit information regarding them. Their supply locally is accomplished by the choice of the type of circuit, the choice of elements to simulate the vocal cords and the eddying constrictions of the vocal system, and the choice of frequency sub-bands. It is now sufficient to transmit information defining the variable characteristics and combine them with the locally supplied fixed circuit features to reproduce the signal.

As will be clear from the foregoing, the system of my prior application, Serial No. 47,393, includes a synthesizer which involves a source of oscillations capable of producing either a discrete frequency spectrum, or voiced sounds or a continuous spectrum for hissing or unvoiced sounds. It also includes a number of control channels in which currents are received for effecting a number of controls. One of these controls performs  
55 two functions. It determines whether the oscillation source will generate a discrete spectrum or a continuous spectrum, and in the former case it determines, in addition, how the fundamental frequency of the discrete spectrum shall vary in  
60 pitch. Other channels are used to control the oscillations thus generated in accordance with the parameters which determine the invariable characteristics of speech, or in accordance with an equivalent set of parameters which may be based,  
65 for example, upon the amounts of energy in chosen sub-bands of the voice range.

In accordance with the present invention, instead of receiving the control currents of these channels from a distant station they may be generated locally by independent manual operated devices such as keys, to produce corresponding modulations and controlling effects upon the locally generated oscillations, thus producing speech manually.

Such an arrangement might be useful in a number of ways. For example, it can be used in certain lines of education or entertainment. It might also find some use as a means for permitting dumb people to talk, or it might be an aid in teaching speech characteristics to dumb people. Various other uses will readily suggest themselves and are within the scope of my invention.

Of course, any device that produces spoken speech synthetically can be used to produce sung speech, that is, vocal music. It has been found, by actual experiment, that the arrangement of my invention will make excellent music, especially where the music is composed particularly for the instrument. In fact, it appears to have extraordinary possibilities to produce music of sorts never heard before. It has been found to make excellent marching music by pressing a large number of fundamental control keys quickly and could be made to produce all sorts of other music by operating the keys in the proper manner. If desired, a number of synthesizers of the type above described may be used for producing chords, as a single fundamental frequency is produced along with its upper harmonics. This corresponds to what occurs in the production of sounds from most musical instruments.

If a number of manually operated speech synthesizers are used in this manner so as to produce a sort of electronic organ, the outstanding characteristics of such an organ will be its ability to copy or simulate vocal music. Many people enjoy music largely because it does imitate the human voice, and there are people who apparently do not enjoy any music except that of the human voice. This type of instrument inherently permits as close an imitation of the human voice as is desired. In this respect it opens up a range between present musical instruments, which tend to be rather mechanical (except possibly in the case of the more complicated pipe organs), and the human voice, which has certain pleasing characteristics, particularly intonation ability and formant or quality-changing ability, which are not found in any other instruments.

Some instruments put out more or less fixed pitches from the notes; whereas, others, such as string instruments, have a range of variation, but put out a more or less fixed quality pattern at the same time, so that they lack considerably as compared to the human voice. These limitations are more or less basic in our present mechanical instruments for producing music and give the music from them a certain mechanical sound characteristic of them. A device of the sort here proposed avoids this difficulty and opens up a huge field of new possibilities in producing music.

The synthetic speech producer of my invention also has the possibility of simulating various instruments. An instrument, such as a violin, typically has a formant, or quality characteristic, due to the resonances of the box of the violin. Similarly with other instruments. By adjusting the amounts of different frequency ranges to be used by the speech synthesizer, different sorts of instruments may be simulated at will.

In addition to its entertainment value, such an arrangement would have a certain amount of educational value along musical lines. This would be particularly true for teaching musical principles and new musical instruments. It would also be useful for developing musical appreciation and, finally, it might be used in voice improvement.

This instrument can also be used for simulating non-musical sounds or a combination of them with music or speech. Thus, it has been used to simulate the departure of a train with first a whistle, then the words "all aboard", then the sound effect "choo choo" starting strong and dying away to a faint rumble. It has also been used to simulate the sound of barnyard animals. These examples are sufficient to indicate the enormous latent possibilities of this instrument for producing a wide variety of sound.

Other objects and aspects of the invention will be apparent from the following description and claims:

Figure 1 shows schematically a system embodying the invention in the specific form referred to above; Fig. 2 is a detail showing a type of finger control which may be used in connection with the invention; Fig. 3 shows a fingering layout of keys, with an arrangement of power control for eight frequency sub-bands by means of the fingers of the two hands, and a pitch control in which the thumbs are employed; Fig. 4 shows a foot pedal control which may be substituted for the thumb control of Fig. 3, thus permitting all ten fingers of the two hands to be used for the power control of frequency bands; Fig. 5 shows the fingering layout of keys for the two hands where the foot pedal is employed for pitch control; and Fig. 6 shows a modified system embodying the invention.

There is disclosed hereinafter the particular circuit which can be used for artificially producing speech or other sound by setting up artificial currents of limited frequency range which simulate the effect of the independent variables of the speech producing organs in man. In my earlier application, Serial No. 47,393, filed October 30, 1935, there is a special analyzing apparatus at the sending end of the system to determine the characteristic of the speech signal being fashioned by the talker. At the receiving end is a synthesizing apparatus to receive these signals and reproduce speech of very close copy, so far as the ear can determine, of the speech at the sending end. In between these two devices there is a transmission line of limited frequency range. The transmission must take place as rapidly as the speech is produced. It is desirable to determine what is the nature of the syllabic or speech determining factors which enable us, by the use of the synthesizer, to control suitable multifrequency generators to produce speech artificially.

In this connection it is convenient to differentiate the production of vowel sounds from those of consonants. As used here, vowels are taken to indicate the pure vowels, the semi-vowels, the diphthongs, and the transitionals. Some thirty-four of these are listed in the book, "Speech Pathology", by Lee Edward Travers. They comprise fourteen vowels, as, for example, a in art; eleven diphthongs as, for example, a in mat; five semi-vowels as, for example, m in me, and four transitionals as, for example, wh in white. These vowel and near-vowel sounds are set forth in detail in my application, Serial No. 47,393.

These vowel and near-vowel sounds have, in general, been produced artificially so that we have a very good check on the number of parameters involved in producing them. From the vocal standpoint, starting at the lungs and coming out to the mouth opening, the variables that may be controlled independently are:

1. Lung pressure.
2. Vocal cords.
3. Rear mouth resonance chamber.
4. Opening from rear to front resonance chamber.
5. Front mouth resonance chamber.
6. Opening from mouth.

These six are more or less in the direct path of speech, but in a sort of by-path we have two more:

7. Nasal resonance chamber.
8. Position of soft palate and uvula (opening to nasal chamber).

The eight of these in their action are not completely independent of one another. Thus, 3, 4 and 5 act decidedly in unison. Some do not, or at least need not, vary greatly, as 6, the mouth opening, which may be kept fixed for the production of all the vowels. Again, the soft palate 8 may open and close the nasal chamber, intermediate positions being unimportant. The eight variables given then, actually may be reduced to five or six in practice.

We come next to the production of the remaining sounds which are classified as fricatives and stop consonants, and are again divided according to whether they are voiced or unvoiced. The voiced ones require the use of the vocal cords; the unvoiced ones do not. As set forth in my application, Serial No. 47,393, above referred to, they comprise eight fricative consonants (four voiced, such as v, and four unvoiced, such as f) and eight stop consonants (four voiced, such as b, and four unvoiced, such as p).

The fricative consonants are produced with about the same position throughout of the vocal organs except that a certain air outlet or aperture is formed at varying places. Thus, for v and f it is formed from the lip to the teeth; for z and s it is formed from the upper teeth to the lower teeth; for the two th sounds it is formed from the tongue to the teeth; for the zh and sh sounds it is formed from the tongue to the hard palate. The voiced consonant is made by pronouncing the unvoiced consonant but vibrating the vocal cords at the same time as though to increase the volume.

The stop consonants are made by forming a stop to the passage of air in the mouth at some particular point, building pressure up behind this and then opening rapidly at the closed point so as to give an explosive sound. The stop is formed by the upper lip against the lower lip in the case of b and p, by the tongue against the upper teeth in the case of d and t, by the tongue against about the middle of the hard palate in the case of j and ch, by the tongue against the soft palate in the case of g and k. In going from the unvoiced to the voiced consonant the formation of the stop, or for that matter, of the opening of the outlet in the case of the fricative consonants, may be slightly further front or backward.

In producing the fricative and stop consonants the different parts of the vocal system are used differently than in the production of vowels. Thus, the nasal resonance is of little importance, the vocal cords are not used in producing the unvoiced consonants, the large air chambers in the front and rear mouth are of much less importance, and two new and very important factors are added, (1) the position at which a closure is partly made and held in the case of the

fricative consonants, and (2) completely made but not held in the case of the stop consonants. To list the independent variables again in the same order as before, we have:

1. Lung pressure.
2. Vocal cords (for voiced consonants).
3. Nasal resonance chamber.
4. Rear mouth resonance chamber.
5. Opening between air chambers of mouth (for fricatives).
6. Front resonance chamber of mouth.
7. Position of closure or explosive opening.

Here, as in the case of vowel production, we have more parameters than are essentially independent with any large degree of freedom. Thus the vocal cords are only used for part of the consonants, the nasal resonance is not very important, the size of the mouth resonance chambers is probably of limited importance. The position of the closure or opening is very important, but the two are essentially the same sort of parameter so they are shown as one rather than two. Accordingly, we conclude again, that of these parameters 5 or 6 are ample to represent the actual variable characteristics in speech production.

There are a number of odd effects that in the discussion up to this point have not been allowed for to any extent, at least, not intentionally. One thing of this sort is odd deformities or deficiencies in the usual oral structure. Other odd effects that we have are those produced when we do unusual things with the voice, such as whisper, talk in a falsetto tone, produce ventriloquistic sounds, or produce what is called double voice.

For all of these odd effects it is probably reasonable to allow two or three degrees of freedom further. However, as the eight original degrees of freedom were considered to be essentially less than eight, it would seem that an allowance of eight for the total might be approximately right. If we desire to be generous, perhaps, we should say ten. In speech over telephone circuits of limited frequency range the number might be 20 per cent or 30 per cent less than that required for high quality speech production; i. e., seven or eight independent variables may suffice for commercial telephone transmission and ten for high quality transmission.

If we vary any of our speech producing variables as rapidly as the controlling muscles permit, we find the limiting speed is about eight or ten times per second. Accordingly, each variable has a fundamental of 10 cycles or less while producing speech.

Having found that there are approximately ten independent variables in speech production, in setting up a circuit for artificial production of high quality speech we then need ten independent parameters. However, we need not use these same ten. So long as the parameters are entirely independent we know mathematically that we can use any ten we choose. Not only can the ten be chosen in any fashion provided they are independent, but if they are not entirely independent, enough more can be chosen to make up for the lack of independence. It is advantageous to pick the ten that from an engineering standpoint give a desirable design or the ten that give an optimum design. A particular case of much interest is that where most of them are the amounts of power in sub-bands of the frequency range of speech.

The interest in this case arises from the fact that it is based on using as parameters those physical quantities that are most easily measured. The easiest thing to measure is power, including current and voltage as measures of power, and the easiest way to separate power into parts for the extra variables needed to measure, is by dividing it up according to frequency bands. After the power frequency characteristic is measured, the sound spectrum to be transmitted is entirely defined except for the power distribution within a frequency sub-band. This last needed factor, in view of the nature of the energy spectrum of speech sounds, is given by the fundamental frequency of the speech sound, considering this frequency to degenerate to zero for unvoiced sounds. Fortunately, we have found means of measuring this fundamental frequency also, thus giving a complete set of specifications for reconstructing the speech sounds.

The frequency pattern in speech seems to be of two types. In vowels and near-vowels there is a fundamental frequency with a large number of upper harmonics. For unvoiced sibilant consonants there is a more nearly continuous energy spectrum somewhat similar (except in amplitude characteristic) to that of resistance noise. For other sounds there may be a mixture of these two patterns with one or the other predominating. For each frequency pattern there is, of course, an amplitude-frequency characteristic.

This dual nature of the speech signals as defined electrically leads to a dual type sending or frequency range reducing circuit and a dual type receiving or speech frequency restoring circuit, as disclosed in my earlier application, Serial No. 47,393, of October 30, 1935. The speech currents entering the analyzer of my said earlier application, energize a frequency pattern control circuit and an amplitude pattern control circuit. The frequency pattern control circuit comprises but one channel and discriminates as to the frequency pattern, that is, as to whether the frequency pattern is a discrete frequency spectrum or a continuous spectrum. This discrimination also includes discrimination as to the fundamental frequency when there is one. The amplitude pattern control circuit branches into ten channels and determines what frequency amplitude pattern we have in each of ten sub-bands of the voice range. The information obtained from these two analyzing elements is expressed in the form of electrical currents whose potentials may be applied to the synthesizer in order that the speech may be reproduced.

Fig. 1 of the present application shows a synthesizing arrangement similar to that employed in my prior application, Serial No. 47,393. This includes a frequency pattern control channel  $FP'$  and a number of amplitude pattern control channels  $AP_1'$  to  $AP_{10}'$ , inclusive. In my prior application, Serial No. 47,393, the signals from the analyzer are applied to the frequency pattern control circuit corresponding to  $FP'$  in Fig. 1, and to the amplitude pattern control channels corresponding to  $AP_1'$  to  $AP_{10}'$  of said Fig. 1. The potential applied to the frequency pattern control circuit  $FP'$  of Fig. 1 is applied across resistances  $B_1$  and  $B_2$  of Fig. 1 to control the frequency pattern sources  $RN$  and  $MV_0$  so as to cause current of the proper frequency pattern to flow from these sources in a manner which will be more fully explained later. The potentials applied to the amplitude control channels  $AP_1'$  to  $AP_{10}'$  of the synthesizer, are used to control

shaping networks  $SN_1$  to  $SN_{10}$  in the respective channels to give the proper amplitude-frequency pattern to the power received from the energy source  $RN$  or from the multivibrator  $MV_0$ , as the case may be.

One possible basis for selecting the frequency bands to use is that of equal importance to articulation. This is a standard commonly used in telephone circuits. As the device shown here may often be used to produce speech directly, a much better standard is perhaps that of noticeability of the absence of different frequency bands. On this basis, a somewhat different set of frequency bands is obtainable, particularly for the lower frequencies. Since the ear is the ultimate observer of a speech sound, the characteristic of the ear is very important in determining what frequency bands to use. It has long been known that the ear observes equally well equal percentage increments of frequency rather than equal increments of frequency. In other words, the ear is essentially of such a nature as to observe the logarithm of frequency rather than frequency directly. This then leads to a general plan in accordance with which the ear is presented with frequency bands having a constant percentage of increment from the lowest frequency to the top frequency, rather than with frequency bands of equal width.

One must not consider the ear, however, to the complete exclusion of the sound producing elements. In the case of the mouth producing speech sounds, it has been found (see section 9, page 17, of *Electrical Engineer's Handbook—Electric Communication and Electronics*, by Pender and McIlwain) that the resonant frequencies of the speech sounds occur more or less uniformly distributed over the frequency range 200 to 6400 when plotted on a logarithmic basis. Accordingly, so far as the generation of speech sound goes, the logarithmic division of the frequency space is satisfactory. In other than speech sounds, such as music, noise, or other sound effects, the fact that the ear observes on a logarithmic basis should be sufficient to insure that this is a satisfactory basis to use.

At very low frequencies, the logarithmic division of frequency space cannot be continued without running into an indefinitely large number of frequency bands. However, it is well known that the sensitivity of the ear drops off greatly at these low frequencies. Accordingly, it has seemed desirable to divide up the low ones, say those below 450 cycles into two bands; zero to 225 cycles per second and 225 to 450 cycles per second. The ends of frequency bands thereafter are at 700, 1000, 1400, 2000, 2700, 3800, 5400 and 7500 cycles to give the ten bands desired. These frequency bands may not always be the best for a particular purpose. This can, in any case, be determined by trial, but since these bands are selected on the basic characteristics of hearing, they are of general utility and well suited for general purposes. A certain amount of deviation from these bands is, of course, permissible without serious detriment.

In the system of my application, Serial No. 47,393, the frequency pattern control circuit must perform a number of functions. At the analyzer it must analyze the speech signal to determine its characteristics with respect to the frequency pattern, that is, it must determine whether the sound is a voiced sound involving a discrete frequency pattern or whether it is an unvoiced sound involving a continuous frequency pattern.

If the pattern is of the former type it will include a fundamental and harmonics thereof, and the fundamental will from time to time vary in pitch so that the harmonics will be raised or lowered in the frequency range as the pitch varies. Consequently, the circuit will also have to determine the pitch.

At the synthesizer the frequency pattern control circuit must determine whether the multivibrator source  $MV_0$  (see Fig. 1) is to be set into operation or whether the resistance noise source RN is to be used, this selection depending, of course, upon whether the analyzed speech sound involves a discrete pattern or a continuous pattern. If the multivibrator source  $MV_0$  is put into operation, it must also be controlled by the frequency pattern control circuit (FP' of Fig. 1) to generate the fundamental corresponding to the fundamental in the speech sound, together with the necessary harmonics.

The operation of the frequency pattern control circuit in its selection, as between a discrete spectrum and a continuous spectrum, takes advantage of the fact that in vowels and other sounds, having a finite fundamental frequency, there is a high power level in the range from 80 to 320 cycles, while in sounds like the unvoiced sibilant consonants, where the power is in a continuous spectrum rather than in a discrete one, the power is much lower. When a speech sound having a high level discrete spectrum condition is to be simulated, the frequency pattern control circuit (FP' of Fig. 1) is energized by a current of such value as to indicate what the fundamental frequency is, without, however, indicating anything about the amplitude of the fundamental frequency in the speech signal. When a low level continuous spectrum speech signal, such as that of a syllabic consonant, is to be simulated, the frequency pattern control circuit FP' is not energized. In the latter case the continuous spectrum pattern generated by the source RN is made available.

The frequency pattern control current which is transmitted to the synthesizer in accordance with the principles of the system disclosed in my application, Serial No. 47,393, is a substantially zero current in the case of a continuous spectrum, but in the case of a discrete spectrum it is of considerable amplitude, and this considerable amplitude varies in accordance with the frequency of the fundamental of the voiced sound. The result is that in the latter case, the amplitude of the frequency pattern control current is able by its variation to determine the fundamental frequency of the discrete frequency pattern that is to be generated at the synthesizer. In the case of a discrete spectrum, therefore, the fluctuating direct current in the frequency pattern control circuit (FP' of Fig. 1), whether it be transmitted from a distant synthesizer, or whether it be generated manually in accordance with the present invention, serves two purposes.

First, it effectively disables the amplifier VA in Fig. 1 which would otherwise amplify the resistance noise received from the resistance R through the amplifier A. The biasing current in the circuit FP' is so applied to a grid biasing resistor  $B_1$  for the amplifier VA, that when substantially no bias is present (as is the case for a continuous spectrum), the resistance noise from R through A is passed on through the amplifier VA. However, when a substantial bias is present, as is the case with the discrete spectrum, the gain of the amplifier VA is decreased by a

negative bias being applied, so that substantially no resistance noise is transmitted.

Second, the current from the circuit FP' is applied to a biasing resistance  $B_2$  in the common grid lead of a push-pull vacuum tube circuit VR. The grid circuits of the two tubes of the amplifier VR, it will be noted, are connected in parallel, but the plates are in series. The purpose here is to control the plate resistances of these tubes by the biasing current. The plate resistances in series are used as the resistance element  $R_0$  of a multivibrator circuit  $MV_0$ , so that the frequency of the multivibrator circuit is controlled by this variable plate resistance  $R_0$ . It is controlled in such a way as to set up the desired fundamental frequency of voice plus all of its harmonics. To insure both even and odd harmonics, the circuit is arranged to take off the output from the two tubes of the multivibrator in series and in parallel, and then combine these two so as to generate all the harmonic frequencies. Another possible arrangement of the multivibrator is to have it designed so as to generate one-half the fundamental frequency from which only the even harmonics are used. With the arrangement as shown, however, the fundamental frequency generated and the harmonics thereof will vary in frequency in accordance with the amplitude of the biasing current, which in turn varies in accordance with the frequency of the fundamental in the voiced signal.

The foregoing applies to the case where the spectrum is discrete. Now let us take up the case of a continuous spectrum. When the signal involves a continuous spectrum, no bias current, or at least substantially no bias current, is present. Under these conditions, the multivibrator circuit  $MV_0$  stops oscillating, and as the amplifier VA is unbiased at the resistor  $B_1$ , the resistance noise is amplified and transmitted.

The multivibrator output and the resistance noise circuit output from the variable gain amplifier VA are combined in the circuit leading to the amplitude controlling circuits through filters  $F_1'$  to  $F_{10}'$ , inclusive. The multivibrator output is first passed through an equalizer  $E_4$  which serves to make the output power the same for each frequency, fundamental and upper harmonics. If desired, this end can be obtained by making the coupling loose between the primary and secondary windings of the multivibrator output transformers, the equalizer  $E_4$ , in this case, being omitted.

In my application, Serial No. 47,393, above referred to, arrangements are provided at the analyzer to determine the variation with frequency of the amplitude of the fundamental frequency components of voiced sounds. To accomplish this result the current at the analyzer, which corresponds to the fundamental frequency component of a given vocal sound, has its syllabic frequency component detected and transmitted to the synthesizer. Such a current, when transmitted to the bias resistor  $B_2$  of a synthesizer such as shown in Fig. 1, determines the fundamental frequency of the multivibrator  $MV_0$ . Consequently, the voltage transmitted from the analyzer should be of such value that the voltage across the resistor  $B_2$  in Fig. 1 will have the proper value to cause the multivibrator to generate the desired fundamental frequency. The fundamental frequency set up by the multivibrator  $MV_0$  will increase and decrease in the same manner as the fundamental frequency of speech sound waves in the analyzer. In the present in-

vention, of course, instead of receiving a control current from a distant analyzer, the frequency pattern control current is generated manually as, for example, by a foot pedal shown in Fig. 1, and is applied to the circuit FP' to control the multivibrator MV<sub>0</sub>, as above described.

The result of all this is that a frequency pattern will be applied to the common circuit leading from the filters F<sub>1</sub>' to F<sub>10</sub>', inclusive, of the synthesizer shown in Fig. 1. This frequency pattern will be continuous and extend over the entire voice range from zero to 7500 cycles in the case of an unvoiced sound. In the case of a voiced sound, the frequency pattern applied to the common circuit of these filters will be a discrete frequency pattern having a fundamental and its harmonics, with the fundamental varying up and down in accordance with the pitch of the voiced sound. These results will be accomplished by so operating the foot pedal 0 in Fig. 1 as to generate no current in the circuit FP' when an unvoiced sound is to be produced, but generate a current of substantial value which varies in the proper manner when a given voiced sound is to be reproduced.

The next matter to be considered is how the frequency patterns thus generated at the synthesizer are to be controlled and modulated to reproduce speech, for it will be clear that unless they are modulated in some manner, we will merely hear a resistance noise sounding somewhat like the roar of the surf at the seashore in case the resistance source is active, and in case the multivibrator source is active, we will merely hear a sound somewhat like that of an ordinary buzzer. In order to modulate these sounds to produce speech, therefore, the amplitude pattern control circuits AP<sub>1</sub>' to AP<sub>10</sub>' are provided and, as herein shown, they may be ten in number, although a lesser number may be used, as will be pointed out later.

In the invention disclosed in my application, Serial No. 47,393, above referred to, amplitude pattern measuring circuits corresponding to the control circuits of the synthesizer of Fig. 1 are provided at the analyzer. These amplitude pattern measuring circuits at the analyzer are essentially circuits which measure how much power there is in the speech signal in a suitable number of chosen small frequency bands, and this information is transmitted by control currents to the synthesizer, where the output of resistance noise from amplifier VA or multivibrator harmonics from the multivibrator MV<sub>0</sub> are shaped accordingly. These frequency bands are chosen as described previously.

For example, at the analyzer of my application, Serial No. 47,393, (assuming it is designed for use with the ten sub-bands already discussed), a speech band in the range between 0 and 225 cycles would be selected from the voice and detected. The detected syllabic frequencies from this sub-band vary in amplitude in accordance with the energy from time to time in this sub-band. Consequently, the detected syllabic current is representative of one of the parameters of speech. Other detected syllabic frequencies from other sub-bands represent other parameters. Taken together, the currents representing these parameters are representative of the amplitude pattern of the vocal sound, and when properly applied at the synthesizer, they modulate and control the frequency patterns generated by the resistance noise source or the multivibrator source, as the case may be.

In the case of the arrangement disclosed in my application, Serial No. 47,393, the variable direct currents representing the amplitude patterns, operate on the synthesizer in the following manner: A variable direct current of syllabic frequency corresponding to the sub-band from the 0 to the 225 cycle range, for example, is applied to a biasing resistor B<sub>3</sub> (see Fig. 1) to give a grid bias to a signal shaping network or push-pull amplifier SN<sub>1</sub>. This bias will vary in accordance with the power of that portion of the speech band which was selected at the synthesizer. The amplifier SN<sub>1</sub>, consequently, amplifies that portion of the frequency pattern, generated by the multivibrator MV<sub>0</sub> or by the resistance noise source VA, which is selected by the 0-225 cycle speech band-pass filter FP<sub>1</sub>'. The modulated output is then fed through a 0-225 cycle speech band-pass filter F<sub>1</sub>'' to the input of the speech amplifier SA, where the outputs from nine other speech band-pass filters (of channels AP<sub>2</sub>' to AP<sub>10</sub>') are combined to give the original speech signal. The speech currents are then transmitted through amplifier SA to the speech receiving output circuit 4.

In the case of the present invention, as shown in Fig. 1, the amplitude pattern control circuits AP<sub>1</sub>' to AP<sub>10</sub>' are controlled by finger keys 1 to 10, inclusive, instead of being controlled by currents transmitted from a distant analyzer. By properly manipulating these keys the generated frequency patterns may be modulated in amplitude in accordance with any desired amplitude pattern which is characteristic of the desired sound to be produced.

In accordance with the present invention, then, artificial speech may be produced by operating the synthesizer manually. This may be accomplished by manipulating the foot pedal 0 of Fig. 1 to control the frequency pattern determining means, and by manipulating the keys 1 to 10, inclusive, with the fingers, for example, to determine the amplitude pattern control.

A manual control adapted to be operated by the fingers is illustrated in Fig. 2, and a corresponding arrangement for foot operation is shown in Fig. 4. These manual controls must be arrangements, each capable of generating a current which increases as the finger pressure or foot pressure is increased. The relation between the output current and the finger pressure may be any relationship found convenient. For the pitch or fundamental frequency it is probably desirable to have the output current more or less proportional to the logarithm of the pressure applied, although an arrangement giving relatively higher frequency than this, at the lower pressures, could also be used readily. For the channels controlling the amount of power, it may be desirable to operate on a logarithmic basis so that equal increments of pressure give equal logarithmic increments of current. In some cases a scale between an arithmetic scale and a logarithmic scale might be desirable. Any type of scale may be chosen, the person using the scale, of course, having to adjust the operation of the manual control device to correspond.

In the typical finger control arrangement of Fig. 2, as the finger pushes the finger-rest F, a rheostat R is adjusted to give the desired current from the battery B. A spring S is provided for restoring the finger-rest when the pressure is removed so that no signal defining current passes when there is no finger pressure applied. The normal or rest condition corresponds to having

the rheostat open, or, in other words, to having an infinite resistance in the circuit of the battery. As a guide for the shaft of the finger-rest, a cylindrical guide is provided with an opening or slot through which the contact finger of the rheostat projects.

The foot pedal arrangement of Fig. 4 is a structure similar to the arrangement of Fig. 2, except that the piston-like member which moves in the guide G is operated by means of a foot pedal P in an obvious manner. As the foot pedal will preferably be used to control the frequency pattern circuit FP', it will, as previously stated, be so designed as to produce a current from the battery B which is more or less proportional arithmetically to the applied pressure.

The method of operating a speech producing circuit such as shown in Fig. 1, by manual manipulation, may be any one which is found convenient. In Fig. 5, for example, is shown a fingering layout of keys for a synthesizer arrangement in which there is one pitch control or frequency pattern control circuit, and ten amplitude pattern control circuits. The five fingers of the two hands are arranged to control the power in the different frequency bands, with the lowest band starting at the left, and the highest one ending at the right, as in the piano keyboard. These controls are represented by the buttons numbered 1 to 10, inclusive, which correspond to the ten sub-bands of the voice. The small left-hand finger then controls the power in the lowest frequency band, and the small right-hand finger controls the power in the highest frequency band. The pitch, or frequency pattern, as shown in Fig. 5, is controlled by a foot pedal P, although it will be obvious that instead of a foot pedal this particular control might be exercised through a suitable mechanism to be held between the teeth or to be manipulated by any other part of the body, as found convenient. The keys may be mounted somewhat, perhaps, like typewriter keys, and in convenient position, so that the hands do not need to move except up and down to apply pressure. Adjustments may be made to get desired positions of keys for different sized hands.

The above arrangement has eleven controls. However, it is not necessary to have as many as ten amplitude pattern controls. These controls may be reduced in number by reducing the voice range somewhat. Thus one or two of the highest frequency sub-bands might be omitted without undue impairment of intelligibility. Again, in some instances, one or two intermediate bands might be omitted without great loss of intelligibility. Another possibility would be to have the frequency range in each sub-band enlarged a bit so that the entire frequency range can be covered by say, eight sub-bands.

By omitting two sub-bands, a fingering layout such as shown in Fig. 3 may be employed. In this figure the eight ordinary fingers of the two hands are arranged to control the power in the different bands with the lowest frequency bands starting at the left and the highest ones ending at the right. The small left-hand finger then controls the power in the lowest frequency band and the small right-hand finger the power in the highest frequency band. The pitch, as shown in this layout, is controlled by a bar to be operated by the thumbs. These keys and the bar would then be mounted somewhat like typewriter keys, and would be arranged in convenient positions for manipulation by the thumbs and fingers. This arrangement, of course, has only nine con-

trols, one for the frequency pattern and pitch control, the others controlling the amplitude pattern in different sub-bands.

The manual controlled system is capable of giving better quality than the arrangement shown in my previous application, Serial No. 47,393, where the synthesizer is controlled by currents resulting from the analysis of actual vocal sounds. This type of system should give better quality than the prior system, so far as distortion goes, since the inherent distortion of the analyzing circuit is eliminated and also since the hands can correct for the distortion, to some extent, if desired. While the number of controls is fairly large, thereby tending to make operation difficult, on the other hand, the amounts of pressure to be applied to the different keys bear certain relations to each other so that the eight or ten keys do not operate entirely independently but, at the most, have one, two or three strongly resonant regions, with the other fingers assuming intermediate positions. Were all nine or ten circuits to control the speech sound produced entirely independently of each other, five or more steps could readily be recognized in each, so that the total number of sounds produced would be  $5^9$  or  $5^{10}$ , an exceedingly large number. Actually, however, there are only forty recognized English sounds. These frequency band controls are used primarily to select these forty sounds. It is therefore fairly simple to operate these if one learns the technique by practicing it for a while. The matter of pitch is somewhat different. It, in a large part, gives the emotional content to speech, so that this single control is required to give this enormous contribution to our speech. In fact, it seems fair to say that one-half of the contribution to speech comes out of pitch which has but a single control. Therefore, the other eight or nine controls should not be much more difficult to learn than is this single pitch control.

While this arrangement for producing speech has been shown as employing the synthesizing arrangement of my previous application, Serial No. 47,393, it is obvious that in principle other arrangements may be used in place of this particular type of synthesizer. The basic principle is that some other portion of the human body than the ordinary vocal organs can produce speech sounds when the proper controls are set up, provided it is required to operate this portion of the body at only its normal muscular frequency range. Therefore, any type of circuit whatever, that gets down to muscular frequencies, that is, frequencies of 5 to 10 cycles per second, could readily be operated to produce speech by hand. It is also obvious that many other simplifications, or extensions, may be provided to a circuit of this sort. Thus, a smaller number of channels or a larger number might be used. The channels might be chosen at different frequency ranges; various portions of the body might be used for controlling; and other modifications will readily suggest themselves.

An instance of such a modification is the circuit shown in Fig. 6. This circuit was developed to make for simpler operation in certain respects than in the case of the circuit shown in Fig. 1. The simplicity results primarily from controlling the output in each channel directly by means of a potentiometer arrangement rather than by going through the intermediate step of setting up a group of control currents from a battery as shown in Fig. 1. This makes for a better modulator in certain respects. It has less equipment, and

works equally well at any level. It can be made to have a large linear range of volume control. There need be no background of noise whatever.

In general, the circuit shown in Fig. 6 approximates more closely to actual voice production than does the circuit shown in Fig. 1 which has been used as the natural development from the analyzer-synthesizer circuit referred to in my previous application, Serial No. 47,393. Other features in which the circuit of Fig. 6 corresponds more nearly to the human voice will be mentioned later.

The elements of the circuit of Fig. 6 consist of a relaxation oscillator such as is described in the copending application of R. R. Riesz, Serial No. 100,291, filed September 11, 1936, a resistance noise source RN, as is shown in Fig. 1 and described in connection therewith, a set of band-pass filters  $F_1'$  to  $F_{10}'$ , as shown in Fig. 1, here illustrated with an external delay equalizer for correcting any delay distortion, a set of finger controls as in Fig. 1 but arranged in a different part of the circuit, a set of bridging resistances to keep the effect of the finger controls confined each to its own channel, two volume controls, an amplifier, and finally, a loud speaker or a telephone line.

As used, the pitch control PC has been operated by depressing it with one of the feet. The energy source selector ESS has been operated either by the twist or side roll of the foot or by means of the wrist. While the energy source selector can be controlled directly from the pitch control as is the case in the circuit of Fig. 1, yet there is some advantage in having these two separate, in that then the exact pitch can be set before the relaxation oscillator is thrown in the circuit by means of the energy source selector switch.

The finger controls 1 to 10 for the amount of energy in the different frequency bands operate as shown in Fig. 1. There is a difference here, however, for in Fig. 1 the absolute volume must be obtained, whereas here, only the relative volume need be obtained. This makes for easier operation of these controls as they can now work over a much more limited range.

The total amount of energy to be produced at any instance is determined by the volume controls. These may be operated by one of the feet or by the bending of the knee, or by other means. The first such volume control is the volume swell VS which will gradually enlarge the volume as the control is depressed. The second such volume control VJ is known as the volume jump because it puts in a sudden change in volume and has been found useful in producing explosive sounds where a sudden change in energy does occur.

It is seen that the finger controls, the volume swell and the volume jump controls all act here to correspond more nearly with what happens in the voice as speech is produced. This circuit corresponds more to the production of the human voice in the arrangement of energy sources. Thus the relaxation oscillator puts out a buzzer-like sound in much the same fashion as the vocal cords do, and does not have equalization such as is shown in the circuit of Fig. 1. Also, the resistance noise can be set to put out the proper relative amount of energy rather than a fixed amount to correspond to the energy from the relaxation oscillator. In these and the other mentioned respects this circuit approaches more nearly the human voice, thereby making it eas-

ier to produce speech sounds as well as insuring the production of more natural speech sounds.

Other modifications may be arranged to make the circuit appropriate to special conditions. For instance, the pitch control can be set so that it naturally operates over the range of an average male voice with a different setting, provided to make the same circuit work over the natural range of an average female voice. Other modifications will readily suggest themselves to one familiar with the art.

What is claimed is:

1. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which consists in manually producing a set of parameters having approximately the number of degrees of freedom of the variable elements of the sound to be produced, producing artificially waves which have a discrete frequency spectrum to represent the invariable information of a voiced sound and which have a continuous frequency spectrum to represent the invariable information of an unvoiced sound, and combining effects of said artificially produced waves and said parameters.

2. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which consists in manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound to be produced, producing artificially waves which have a discrete frequency spectrum to represent the invariable information of a voiced sound and which have a continuous frequency spectrum to represent the invariable information of an unvoiced sound, and combining effects of said artificially produced waves and said defining waves.

3. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which consists in manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound to be produced, the frequency of the defining waves being below audibility, producing artificially waves which have a discrete frequency spectrum to represent the invariable information of a voiced sound and which have a continuous frequency spectrum to represent the invariable information of an unvoiced sound, and combining effects of said artificially produced waves and said defining waves.

4. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which comprises producing a frequency pattern which has a discrete spectrum corresponding to the invariable information of a voiced sound to be produced and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound to be produced, manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound, and modifying said frequency pattern in accordance with said defining waves.

5. The method of producing vocal and other

sounds containing variable information and invariable information and represented by a complex wave, which comprises producing a frequency pattern which has a discrete spectrum corresponding to the invariable information of a voiced sound to be produced and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound to be produced, manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound, the frequency of the defining waves being below audibility, and modifying said frequency pattern in accordance with said defining waves.

6. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which comprises producing a frequency pattern which has a discrete spectrum corresponding to the invariable information of a voiced sound to be produced and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound to be produced, manually producing a set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, and modifying said frequency pattern in accordance with said defining waves.

7. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which comprises producing a frequency pattern which has a discrete spectrum corresponding to the invariable information of a voiced sound to be produced and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound to be produced, manually producing a set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, the frequency of the defining waves being below audibility, and modifying said frequency pattern in accordance with said defining waves.

8. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which consists in manually producing a frequency pattern controlling wave, determining by said controlling wave the production of a frequency pattern having either a continuous spectrum or a discrete spectrum, controlling by said controlling wave the fundamental frequency of the discrete frequency pattern, manually producing a set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, and modifying said frequency patterns in accordance with said defining waves.

9. The method of producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, which consists in manually producing a frequency pattern controlling wave, determining by said controlling wave the production of a frequency pattern having either a continuous spectrum or a discrete spectrum, controlling by said controlling wave the fundamental frequency of the discrete frequency pattern, manually producing a set of defining waves each of which

defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, the frequency of the defining waves being below audibility, and modifying said frequency patterns in accordance with said defining waves.

10. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for manually producing a set of parameters having approximately the number of degrees of freedom of the variable elements of the sound to be produced, means for artificially producing waves which have a discrete frequency spectrum representing the invariable information of a voiced sound and which have a continuous frequency spectrum representing the invariable information of an unvoiced sound, and means for combining effects of said artificially produced waves and said parameters to produce the desired sound.

11. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound to be produced, means for artificially producing waves which have a discrete frequency spectrum representing the invariable information of a voiced sound and which have a continuous frequency spectrum representing the invariable information of an unvoiced sound, and means for combining the effects of said artificially produced waves and said defining waves to produce the desired sound.

12. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound to be produced, the frequency of the defining waves being below audibility, means for artificially producing waves which have a discrete frequency spectrum representing the invariable information of a voiced sound and which have a continuous frequency spectrum representing the invariable information of an unvoiced sound, and means for combining the effects of said artificially produced waves and said defining waves to produce the desired sound.

13. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for producing a frequency pattern which has a discrete frequency spectrum corresponding to the invariable information of a voiced sound and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound, means for manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound, and means for modifying said frequency pattern in accordance with said defining waves.

14. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a com-

- plex wave, means for producing a frequency pattern which has a discrete frequency spectrum corresponding to the invariable information of a voiced sound and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound, means for manually producing a set of defining waves that respectively define the variations of a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the sound, the frequency of the defining waves being below audibility, and means for modifying said frequency pattern in accordance with said defining waves.
15. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for producing a frequency pattern which has a discrete frequency spectrum corresponding to the invariable information of a voiced sound and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound, means for manually producing a set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, and means for modifying said frequency pattern in accordance with said defining waves.
16. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for producing a frequency pattern which has a discrete frequency spectrum corresponding to the invariable information of a voiced sound and which has a continuous frequency spectrum corresponding to the invariable information of an unvoiced sound, means for manually producing a set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, the frequency of the defining waves being below audibility, and means for modifying said frequency pattern in accordance with said defining waves.
17. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for manually producing a frequency pattern controlling wave, means to produce a frequency pattern under the control of said controlling wave having a continuous spectrum under certain conditions of said controlling wave and a discrete spectrum under other conditions of said controlling wave, means controlled by said controlling wave for controlling the fundamental frequency of said discrete frequency pattern, means for manually producing a set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, and means for modifying said frequency patterns in accordance with said defining waves.
18. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, means for manually producing a frequency pattern controlling wave, means to produce a frequency pattern under the control of said controlling wave having a continuous spectrum under certain conditions of said controlling wave and a discrete spectrum under other conditions of said controlling wave, means controlled by said controlling wave for controlling the fundamental frequency of said discrete frequency pattern, a set of manually operated keys for individually producing a corresponding set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, the frequency of said defining waves being below audibility, and means for modifying said frequency patterns in accordance with said sets of defining waves.
19. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, a manually operated key, means controlled by said key for producing a frequency pattern controlling wave, means controlled by said controlling wave under certain conditions thereof to produce a frequency pattern having a continuous spectrum, means controlled by said controlling wave under other conditions thereof to produce a frequency pattern having a discrete spectrum, means controlled by said controlling wave for controlling the fundamental frequency of said discrete frequency pattern, a set of manually operated keys for individually producing a corresponding set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, and means for modifying said frequency patterns in accordance with said set of defining waves.
20. In a mechanism for producing vocal and other sounds containing variable information and invariable information and represented by a complex wave, a manually operated key, means controlled by said key for producing a frequency pattern controlling wave, means controlled by said controlling wave under certain conditions thereof to produce a frequency pattern having a continuous spectrum, means controlled by said controlling wave under other conditions thereof to produce a frequency pattern having a discrete spectrum, means controlled by said controlling wave for controlling the fundamental frequency of said discrete frequency pattern, a set of manually operated keys for individually producing a corresponding set of defining waves each of which defines the variations in amplitude of a separate sub-band of the band of frequencies comprising the sound to be produced, the frequency of said defining waves being below audibility, and means for modifying said frequency patterns in accordance with said sets of defining waves.

HOMER W. DUDLEY.