

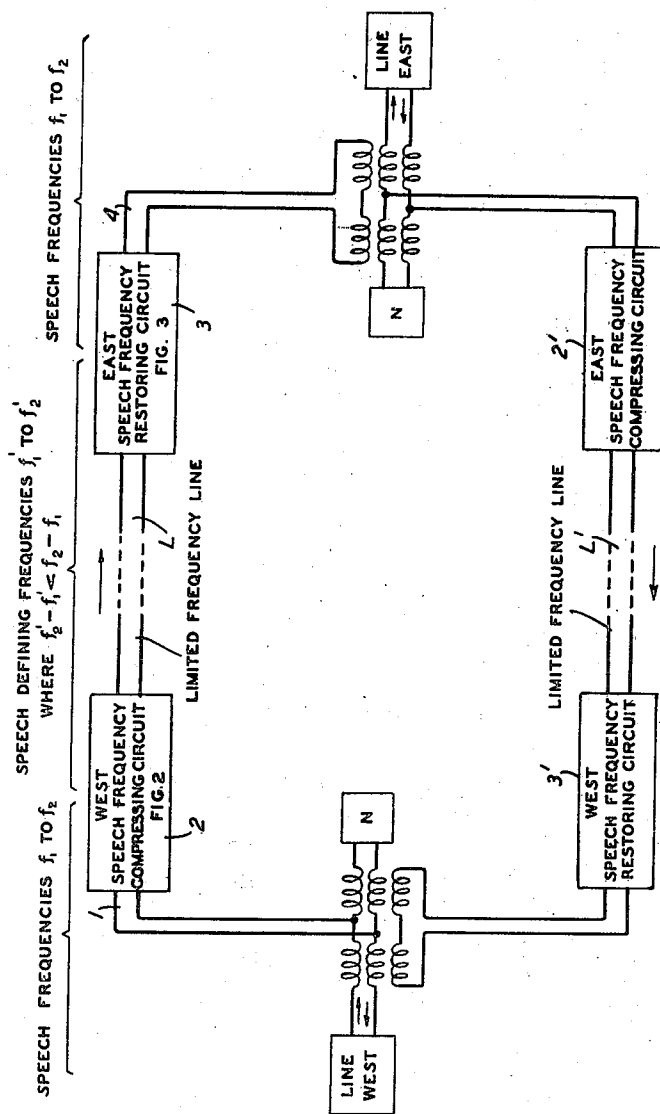
March 21, 1939.

H. W. DUDLEY
SIGNAL TRANSMISSION
Filed Oct. 30, 1935

2,151,091

3 Sheets-Sheet 1

FIG. 1



INVENTOR
H. W. DUDLEY
BY *J. A. Burgess*
ATTORNEY

March 21, 1939.

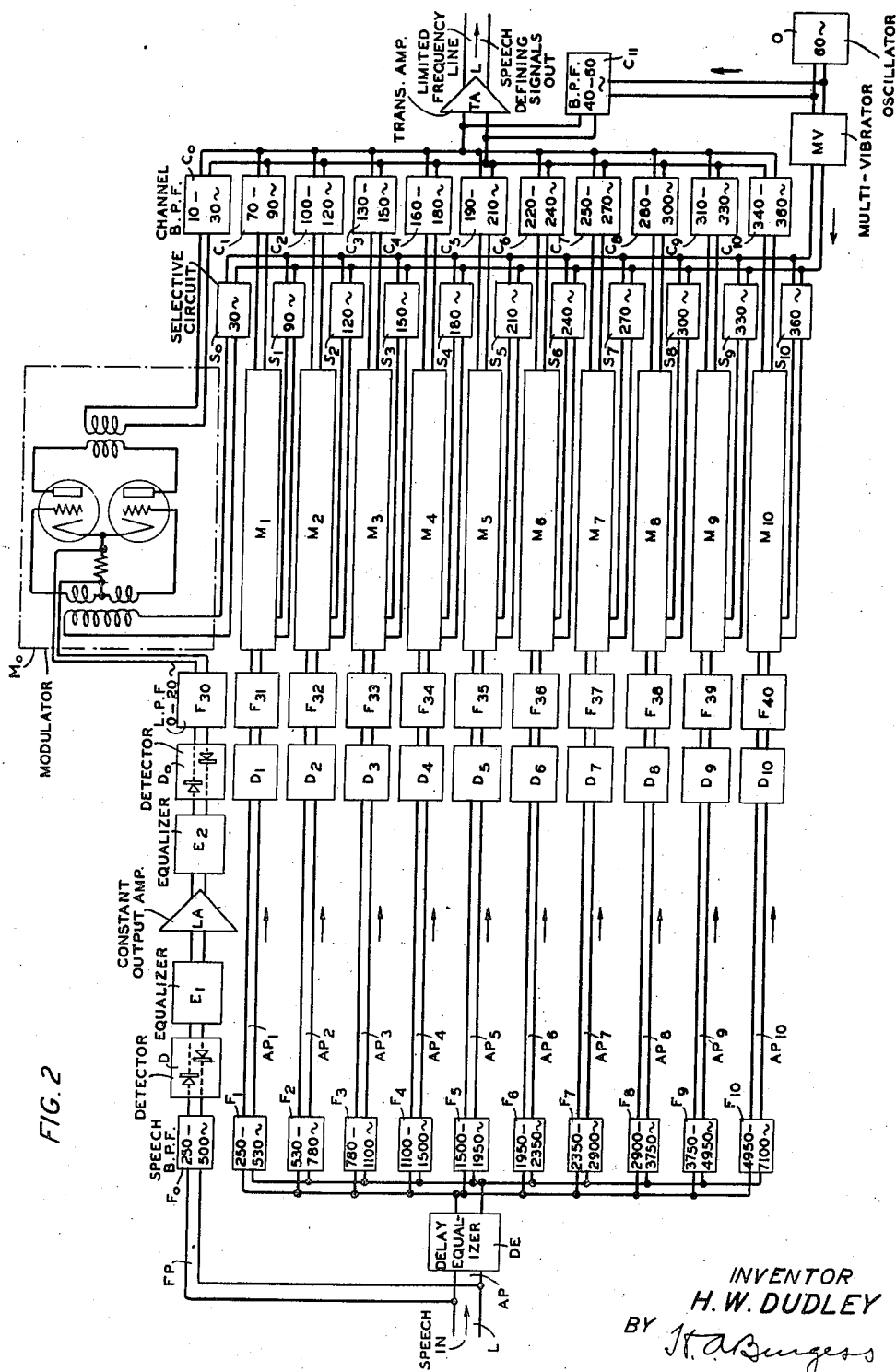
H. W. DUDLEY

2,151,091

SIGNAL TRANSMISSION

Filed Oct. 30, 1935

3 Sheets-Sheet 2



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

March 21, 1939.

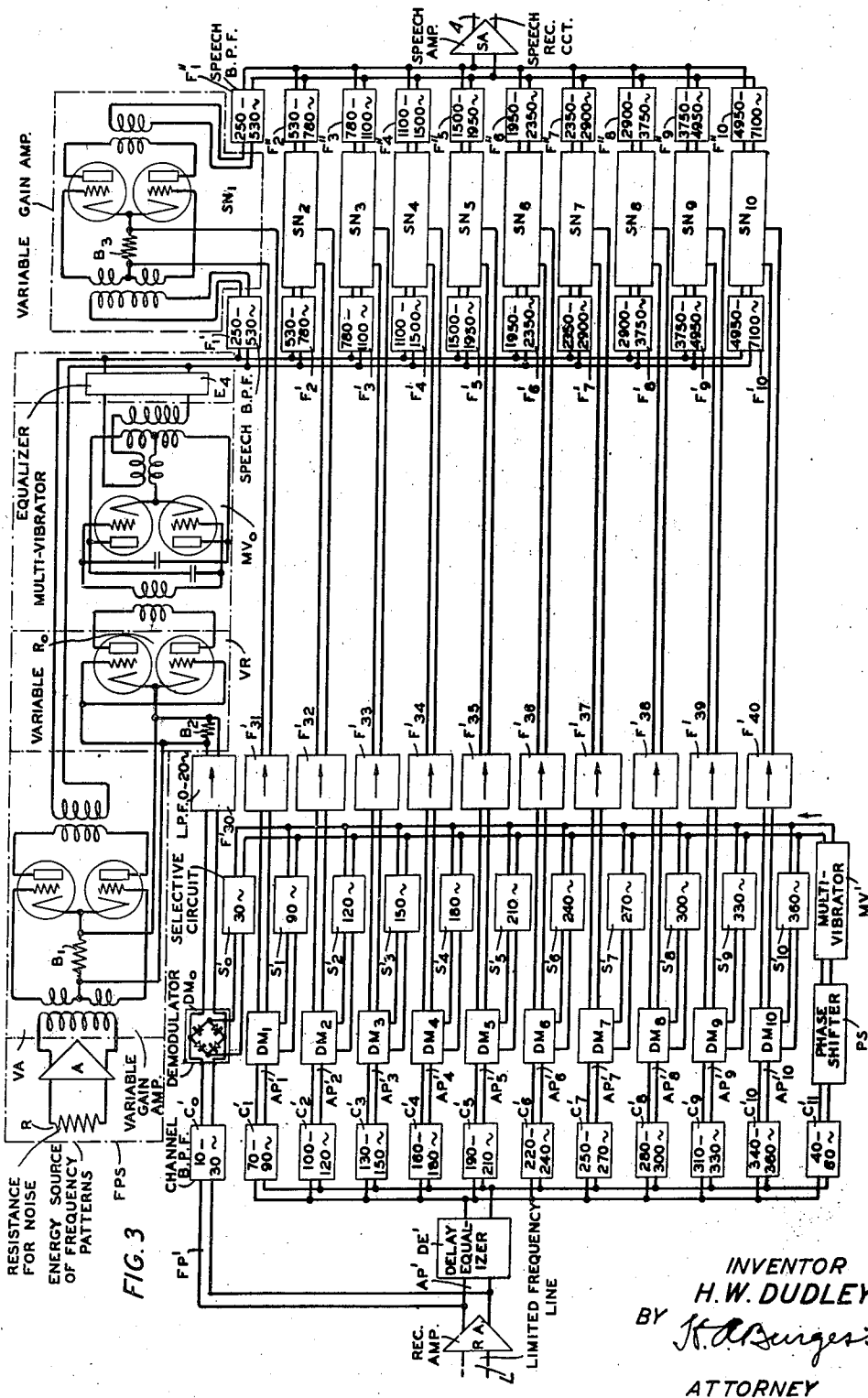
H. W. DUDLEY

2,151,091

SIGNAL TRANSMISSION

Filed Oct. 30, 1935

3 Sheets-Sheet 3



UNITED STATES PATENT OFFICE

2,151,091

SIGNAL TRANSMISSION

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

Application October 30, 1935, Serial No. 47,393

74 Claims. (Cl. 179—1.5)

This invention relates to wave transmission, as for example, signaling.

An object of the invention is to reduce the frequency range required for the transmission of signals, as for instance, speech.

Further, it is an object of the invention to effect such reduction without necessitating increase in the time of transmission.

It is also an object of the invention to facilitate transmission of a message over a transmission medium not adapted to readily pass the frequency band originally occupied by the message, as for example, to facilitate telephoning over long submarine cables where the higher speech frequencies cannot readily be transmitted because of excessive attenuation.

Another object is to facilitate increasing the number of signaling channels that a wave transmission medium, as for example, a land line or a radio circuit, can transmit in a given frequency range.

Another object is to facilitate increasing the signal-to-noise ratio obtainable at the receiving end of a system without increasing the sending level, as for instance, in long distance radio communication where the transmitting power required tends to become excessive.

It is also an object of the invention to promote secrecy in transmission of signals, for instance, in wire or radio transmission of speech.

Objects of the invention are also analysis, synthesis, production, reproduction, reconstruction, simulation, imitation, modification and control of sounds and waves representing sounds, especially speech sounds and vocal effects of usual and unusual character.

The information transmitted by speech does not absolutely require all the frequency space allotted to it in the human voice. A specific case can be worked out as to how much frequency band is required as a minimum, for example, by determining and taking account of the number of the independent variables or parameters involved in speech production (i. e., the number of the independently movable physical elements of the vocal system that are involved in speech production) and the rate at which these can vary. This is done hereinafter, from which it is estimated that speech from 250 to 7100 cycles can be represented almost perfectly by the informa-

tion that can be passed in a 500 cycle band or so using the same time of transmission as the speech itself requires.

The well-known rule of conservation of the product of transmitted frequency band width and time of transmission, formulated by Hartley (Transmission of Information—Bell System Technical Journal, July, 1928), which says that to produce a replica of a given piece of transmission information, the frequency band in the transmitting medium can be decreased by any desired factor only when a time of transmission is used that is longer by the same factor, applies to the case when no information as to the signal source at the transmitting end of the system is available at the receiving end or when such information is available but cannot be used in a practical way at the receiving end to reproduce information which is not transmitted over the line. Also this rule is stated for the case of distortionless transmission. Where there is distortion in the transmission medium some types are less objectionable to the ear than other types, phase distortion, for instance, within wide limits being unimportant in telephony. Thus a given frequency band with a certain amount of distortion of one type may transmit intelligence better than a wider frequency band having distortion of a more objectionable type.

With this understanding as to the rule, it will be clear from the disclosure herein that nothing is proposed herein which is at all contrary to the rule.

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 as considered here are 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, the 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 noted above.

The whole vocal system may be likened to a mechanical-acoustical oscillator with certain fixed circuits and certain variable mechanical elements. This circuit is the same from man to man. It is the same from sound to sound in the same man with the different circuit 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 feed-back may also be controlled. This vocal circuit, fixed in that no different circuit may be substituted therefor, is probably the most important fixed part or element of the vocal system.

Just as the oscillator when oscillating is still essentially a fixed piece of apparatus to which variable controls of frequency, output and feed-back 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 the generated signal 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 100 to 150 cycles per second whereas the variable controls specified can change only at rates ten or more times smaller than this. Strictly the fixed features include 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 constrictions or restricted openings such as between lip and teeth for the "f" sound, between tongue and hard palate for the "sh" sound between the vocal cords themselves for whispering, etc.

From another point of view the variable parts are those controlled volitionally and the fixed parts those controlled non-volitionally if changed at all. Thus one chooses the lip position he desires for a particular sound but in the case of the vocal cords he blows a puff of air through them when they are held with a certain tenseness and they then vibrate at their natural frequency.

This differentiation between fixed and variable features may be applied to characterize the type of signal produced just as it has been applied to the type of element producing the signal. 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 opening, 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. A particular set of variations leads to silence as an extreme case of modulating the normal speech sound. In Harvey Fletcher's book "Speech and Hearing", 1929, there are shown a number of

speech sound oscillograms on pages 33 to 44 which illustrate the characteristics described. As mentioned hereinbefore the fixed features as appearing in a speech signal are oscillatory in nature and the variable features are modulatory. It is seen 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 can be found as an almost identical copy. If the slight change that occurs from period to period is followed up until it totals enough so that the original wave form is lost, this is found to occur many periods later, oftentimes twenty or more requiring .200 second or so. This latter type of change is the modulatory type of change. It is easiest seen in a single sound as a 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 syllable and such change is therefore known as a 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 high frequency patterns. On the other hand we have varying parts or elements setting up modulatory waves of low syllabic frequency pattern.

An ideal frequency range compressing system would be one that took advantage of the dual nature of speech by setting up at the receiving end all the fixed features and transmitting over the line from the sending end only information as to the instantaneous positions of the variable parts. The frequency range compressing system need not be so identical a copy of the vocal system however. 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 mentioned does generally lead to a larger number of required parameters because the newly selected ones are not independent. However, this is not of much practical importance, for the great saving of frequency range accomplished by the invention is obtained by transmitting syllabic frequencies corresponding to the variable or modulatory controls of the speech signal and not transmitting the high frequency pattern corresponding to the fixed oscillatory characteristics of the speech sounds. As this saving is of the order of 10 to 1, a 50 to 100 per cent increase in the parameters required to satisfactorily represent speech is unimportant.

As pointed out in detail hereinafter, the number of the independent variables involved in the production of speech is small. That is, the number of the 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, or 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 these physical or physiological elements, the minimum time in which it can go through a complete cycle of change in position is not less than about a tenth of a second, or in other words, each independent variable has a fundamental frequency of not over

10 cycles per second while engaged in speech production. Further, as explained hereinafter, a frequency band width which need not exceed two or three times the product of this frequency and the number of the independently variable physical elements can suffice to transmit signals containing practically precise information as to the instantaneous displacement values or positions of the independently variable physical elements, and therefore containing the variant or unpredictable information in the speech, and so constituting speech defining signals that define the variable characteristics of the speech. As explained before, these defining signals can be any signals derived from speech signals provided the derived signals give as many independently variable quantities or factors or parameters as the number of independent variables involved in the production of speech; and the chosen parameters need not be entirely independent provided their number be increased sufficiently to make up for their lack of independence. For example, the chosen parameters may be the average amounts of power in subbands of the frequency band of speech waves, as brought out in detail hereinafter.

There exists, then, in the actual production of a complex speech wave by the vocal system, a simple set of slowly varying elements or parameters, (the independently 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 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 number of degrees of freedom of variation.

Like considerations hold in the case of other complex signals, as for example, signals which as in the case of speech, have a sluggish rate of change modulating or controlling a rapid rate of change of signal strength, or from the point of view of frequency, a set of low frequencies modulating a set of high frequencies. Music may be mentioned specifically as one such example. Each note has a fundamental and upper harmonics of a resonant system formed for example from a stretched string, vibrating reed or air tube or chamber, and these fundamental and harmonic components of the notes form the set of high frequencies in the complex signal wave, the low frequency set corresponding to the rates of energizing such resonant systems, as for example by the hand striking the piano keys. Each key is an independently variable element in this case if one considers the keyboard as the starting point. If the hands are considered as the starting point, then the maximum number of fingers used at any one instant are the independently variable parameters. The rest of the piano mechanism is of a fixed nature.

In accordance with one feature of the invention, to transmit a complex signal, for example speech, variant information in the signal or information regarding the variable or unpredictable characteristics of the signal is transmitted, instead of the complex signal wave. The waves transmitted can define the signal precisely, as regards its unknown or variable characteristics, yet have small frequency range relative to the

signal wave and as short duration as the signal. The signal defining waves transmitted need not directly contain any intelligibility whatsoever. Thus a high degree of secrecy is attainable.

In accordance with a feature of the invention as applied to transmission of speech, for example, advantage is taken of the fact that much of the information ordinarily transmitted is of an invariable or predictable character, due to the general uniformity of the speech producing organs from person to person, by reproducing such predictable information artificially at the receiving end of the transmission system in order that it need not be transmitted. Thus, effective use is made of 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 is a system in which a speech signal is analyzed for its fundamental frequency and the average power in properly chosen subbands of frequency, this information being transmitted either directly or by carrier currents, for example, and then used at the receiving end to fashion waves from a local multi-frequency source into a simulation of the signal.

To fashion the simulation of the signal from the waves supplied by the local source, frequency subbands of those waves are selected which are respectively coextensive with the chosen subbands of the speech signal, and the average power in each subband of the locally supplied waves is varied in accordance with the power in the corresponding chosen subband of the signal wave.

This variation is effected in response to the information transmitted from the sending end of the system regarding the average power in chosen subbands of the signal wave.

As indicated above, two types of power spectra or frequency patterns of energy are used alternately in speech, (1) a continuous spectrum in the case of hissing or unvoiced sounds and (2) in the case of voiced sounds a discrete spectrum with a variable fundamental and with all of the upper harmonics always present. Hence, the local source provided is preferably such that the waves supplied by it can have either type of power spectrum. The type and the fundamental frequency can be varied in response to the information transmitted from the sending end of the system regarding the presence or absence of a fundamental frequency in the speech wave and the magnitude of any such fundamental frequency.

Since 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 the muscles of the vocal system to less than about 10 cycles per second, (a frequency much lower than the fundamental frequencies, such as 60 to 500 cycles, of vocal cords), it results that the equipment required at the sending end of the system for analyzing the speech signal for fundamental, and likewise the equipment provided at 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 10 cycles per second, just mentioned, depending on the accuracy desired in the transmission of the indications.

Moreover, the number of subbands analyzed for

power content need not exceed five or ten for example, to obtain high intelligibility; because 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 subband is largely independent of that in the others).

Such a system, then, analyzes the signal into fixed features and variable features. The fixed features include (a) the existence of definite frequency subbands in which the power distribution is sensibly uniform, (b) the existence of a frequency spectrum that alternates from the continuous type to a discrete type with varying fundamental and with all upper harmonics always present, and (c) the fact that time variations of the fundamental frequency and of the power in the frequency subbands occur only at syllabic frequency rates. The variable features include (A) the magnitude of the average power in each subband, 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 subbands. 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 indicated above and hereinafter, the invention is by no means limited to this specific form.

In accordance with a feature of the invention the character or expression or vocal effect of synthesized or reproduced vocal sounds can be controlled or modified by suitable control of their fundamental frequency.

For example in the frequency range reducing and restoring system described above, when normal reproduction of speech is desired the variation of the fundamental frequency in the reproduced speech is made the same as in the original speech; or in other words, when normal reproduction of speech is desired the value of the ratio between the fundamental frequency set up at the reproducing end of the system and the fundamental frequency of the input at the sending or analyzing end is maintained at unity. Further, the value of the ratio may be made to undergo changes or deviations from unity, or the fundamental frequency of the speech or vocal sounds may be made to undergo modifications, suitable for producing prescribed modifications in the speech or vocal sound effects. The modifications may, for example, be such as to give rise to odd voice effects which simulate effects commonly observed in human voices. For instance, a person can talk in an ordinary voice and yet give rise to speech that has the character of a whisper or the character of a chant, for example, dependent upon the condition or adjustment of the system. With the system in proper condition or adjustment he can get the prescribed effect with the uniformity of machine production. Thus a single person, though talking in a normal voice can produce different voices of prescribed types, by having the system

in different conditions or adjustments. A whisper can be obtained by manufacturing the speech or voice from energy having only a continuous spectrum, for example, resistance noise. (Energy having a continuous frequency spectrum or frequency pattern can be viewed as energy with its fundamental frequency infinitesimal or substantially zero, or in other words, can be viewed practically as substantially energy with no fundamental frequency present.) Thus, the whisper can be obtained by conditioning or adjusting the system so that the local source at the receiving end of the system, which supplies the waves from which the speech is reconstructed, will supply waves of only one type of spectrum, the continuous spectrum, (instead of supplying waves of both the continuous type of spectrum and the discrete type of spectrum as when natural reproduction is desired). Then sounds voiced in the original speech are reproduced as unvoiced sounds. A chant can be obtained by manufacturing the voice from energy having only a discrete type of spectrum and a fixed fundamental frequency, as for example, by conditioning or adjusting the system so that the local source at the receiving end of the system, which supplies the waves from which the voice is reconstructed or manufactured, will supply waves of only the discrete type of spectrum and with their fundamental frequency set at an arbitrary chosen frequency such as the average vocal cord frequency instead of variable with the fundamental frequency in the original speech as when natural reproduction is desired.

It is seen, then, that in accordance with the invention, speech can be manufactured or artificially created from different types of energy spectra, as for example, from energy having a continuous frequency spectrum or from energy having a discrete frequency spectrum, or from both types of energy, by relatively varying the time flow of energy in frequency subbands of the spectrum substantially in accordance with the relative variations of the time flow of energy that are characteristic of respectively corresponding subbands of the frequency components of speech sounds to be created, for instance, speech sounds that have been analyzed with a view to their reproduction or reconstruction; and the character or expression or vocal effect of the specific sounds created depends on the fundamental frequency of the energy spectrum used to create them, for example, the manufactured speech having the monotone of a chant when the fundamental frequency is maintained at a fixed value such as the average vocal cord frequency, or having the sound of a whisper when the fundamental frequency approaches zero, i. e., when the energy spectrum is continuous.

A specific aspect of the invention is artificial production, manufacture or synthesis of speech of prescribed character, (for example, speech of usual character, or speech of unusual character such for instance as a whisper or a chant), with the aid of information as to the frequency pattern and the amplitude pattern, or in other words, information as to the three characteristics, pitch, resonant frequency regions and loudness, of the speech sounds desired.

Also, a specific aspect of the invention is determining or obtaining these three characteristics for any particular set of speech signals by instantaneous analysis of the speech signals.

Also, a specific aspect of the invention is synthesis or artificial reconstruction or reproduction

of speech signals in accordance with such analysis.

Further, a specific aspect of the invention is reconstruction of speech with prescribed modification in character or vocal effects.

Also, a specific aspect of the invention is controlling the character of synthesized or reconstructed speech by artificially modifying the fundamental period of speech relatively to the speech duration.

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

Fig. 1 shows schematically a system embodying the invention in its specific form referred to above;

Fig. 2 is a detailed showing of the transmitting circuit at the west terminal of the system, this transmitting circuit being the west speech frequency compressing circuit shown as the block in the upper left-hand corner of Fig. 1; and

Fig. 3 is a detailed showing of the receiving circuit at the east terminal of the system, this receiving circuit being the east speech frequency restoring circuit shown as the block in the upper right-hand corner of Fig. 1.

There is disclosed hereinafter a particular circuit which can be used for reducing the frequency range of speech by setting up artificial currents of limited frequency range which will transmit information about the independent variables of the speech producing organs in man. The first important question is: What frequency range does speech inherently require for transmitting its information content?

In the operation of such a system with high quality transmission of speech, a person at one end of a telephone line talks into a high quality subset. At the distant end the listener also has a high quality subset. The transmission between the two is not direct but is carried on in the following way. At the sending end is special apparatus to tell the characteristics of the speech signal being fashioned by the talker. At the receiving end is other apparatus to receive these signals (transmitted directly or transmitted indirectly as by carrier currents) and produce speech that is a very close copy so far as the ear can determine of that at the sending end. In between 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 roughly what is the minimum frequency band that can be used on the transmission line for such high quality transmission.

A rough answer to this problem can be obtained by various methods. One method is used hereinafter to obtain an answer and another to check that the answer is fairly conservative. The method used for obtaining an answer is essentially as follows. We determine how many independent variables are involved in the production of speech. Next we determine the minimum amount of time in which one of these variable elements can go through a complete cyclic change of position. From these the frequency band required for the transmission of reasonably precise information as to the instantaneous position of each independent physical variable is obtained. Next multiplying this frequency band by the number of independent variables we have the total minimum frequency band required. For actual transmission in circuits allowances

must be made for the inefficiency of filter and other auxiliary circuits that will be required.

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 Mr. Lee Edward Travers. They are as follows:

14 VOWELS

art	moth
fast	law
debris	obey
set	put
fur	truth
it	above
see	up

11 DIPHTHONGS

mat	poor
care	mile
cafe	mouse
here	oyster
so	bare
bore	

5 SEMI-VOWELS

me
no
ring
lady
ride

4 TRANSITIONALS

white
yes
he
we

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 circuit of speech whereas for a 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 do not act 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 or close the nasal chamber, intermediate positions being unimportant. The eight variables given then probably reduce 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. They are:

8 FRICATIVE CONSONANTS

5	4 Voiced	4 Unvoiced
	<u>y</u> ine	<u>f</u> ight
	<u>z</u> ero	<u>s</u> ee
	<u>t</u> hese	<u>t</u> humb
	<u>a</u> zure	<u>sh</u> e

10 8 STOP CONSONANTS

15	4 Voiced	4 Unvoiced
	<u>b</u> e	<u>p</u> ail
	<u>d</u> o	<u>t</u> o
	<u>j</u> oy	<u>ch</u> urch
	<u>g</u> o	<u>k</u> ite

The fricative consonants are produced with about the same position throughout of the vocal organs except that a certain air outlet or small 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 in the position at which a closure is partly made and held in the case of the fricative consonants and 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 as 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 have not in the discussion so far 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. These may be in the form of missing teeth or teeth with separation, odd shapes of various parts of the palate, tongue, etc. Another type of 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. Whispering probably does not differ appreciably from other sound production so far as a typical analyzer of voice is concerned. It is apparently produced by closing off the major portion of the vocal cord opening (glottis) by some muscles which originate and end in the vocal cords themselves. This leaves a small opening at the rear of the vocal cords which produces two predominant effects:

1. Low volume since only a limited amount of air can get through.
2. More high frequencies since the short span tends towards high vibration rates with eddy currents rather than a natural frequency.

However, an analyzing circuit presumably will take care of frequency as a single thing so that provided it operates over a wide enough frequency range nothing special is required. The same is true of volume or power output. For all these odd effects it is probably reasonable to allow two or even 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 well be approximately right. If we want to be generous perhaps we should say ten. In speech over telephone circuits of limited frequency range the number might well be 20 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.

We can get a rough check on whether these eight or ten degrees of freedom are sufficient by considering the variety of distinguishable sounds we would hope to get with this many degrees of freedom as compared to the number we might compute for actual speech. For actual speech the number of sounds for telephonic purpose is generally taken as around fifty such as the fifty given above. Perhaps the most that need be recognized for our purpose is around one hundred. [The International Phonetic Alphabet contains seventy-six sounds, (no diphthongs included) from a considerable variety of languages. In addition five markings are used to indicate (1) length, (2) stress, (3) tenseness or laxness of position, (4) syllabic consonants and closing consonantal parts of diphthongs and (5) nasalized vowels. Only the last seems to offer a distinct sound type for the computation being made here. Thus length applies to the time the sound is being given and not to its type; stress refers chiefly to the instantaneous volume; tenseness or laxness of position affects the pitch and voi-

ume; while syllabic consonants and the closing consonantal parts of diphthongs and for that matter diphthongs themselves represent distinct sounds enunciated in rapid succession. The nasalized vowels like the French *a* do represent distinct sounds for our classification but of the twenty-seven vowel sounds many are not nasalized. Accordingly a total of one-hundred distinct sounds i. e. the seventy-six listed plus twenty-four for the nasalized vowels seems not unreasonable to use here.] The same sound can be given at different volume levels. Probably a 30 decibel range with sixty recognizable steps of $\frac{1}{2}$ decibel each would correspond to a very fine gradation basis. Fifteen steps of 2 decibel would correspond more nearly to ordinary telephone practice. As to pitch, a 1 per cent change is probably noticeable on a fine basis which would lead to about seventy steps per octave or one hundred forty steps altogether with two octaves. On a cruder basis corresponding more nearly to telephone practice probably twenty steps would be plenty. The only remaining item might be classified as personality. We can recognize a certain individuality or personality of sounds so that with a number of sounds from a certain person we succeed in guessing who is talking provided we are familiar enough with the person's voice. It is a little hard to say what is the maximum number of people we might hope to distinguish when a single sound is pronounced at a given pitch and volume. In the first case we will generously assume it may be as great as five hundred. More practically we will assume it is likely not to exceed ten. These sets of figures give us for recognizable elements in the fine grained picture a total number of

$$100 \times 60 \times 140 \times 500$$

or about 4×10^8 . In the coarser picture we have $50 \times 15 \times 20 \times 10$ or about 15×10^4 .

Coming now to our artificial speech production system, the number of steps we can recognize in each independent variable multiplied by the number of such variables gives us the total number of combinations available. The observer that recognizes the changes in the variables is essentially our receiving circuit rather than the ear so that the number to put down is the number corresponding to the ability to differentiate steps of signal level in our chosen machine rather than in the human ear. Probably to recognize twenty gradations is not unduly optimistic. To recognize ten is reasonably conservative. On the basis of recognizing ten and having ten parameters we conclude we should recognize altogether 10^{10} different sound elements which is a little greater than the largest number we figure that we may be required to recognize. Accordingly the ten parameters seem to be quite ample from this rough check.

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

We are now in a position to calculate the band width required for transmitting the speech defining signals. On the basis of transmitting as the useful band not more than 100 per cent of the fundamental frequency (this is done in television and telephotography) we require a total frequency band of $10 \times 10 \times 1.00$ or 100 cycles for transmitting the speech defining signals. This

100 cycles is more or less of an absolute minimum for high quality speech no matter how much money we are willing to spend. In telegraph more than 100 per cent of the fundamental frequency is required as the minimum transmitted band. Thus, in non-loaded submarine telegraph cables it is about 110 per cent (H. Nyquist); while in voice frequency telegraph where frequency space is less important it uses about 350 per cent for either side-band and transmits both. A 200 per cent basis is quite conservative. This would mean 200 cycles of useful frequency band.

A convenient method of transmitting such speech defining information is by a miniature carrier system of ten channels, one for each independent variable. In this case filters are required to separate the channels. These filters cannot cut off sharply but require waste space between used frequency bands. This waste frequency space may amount to 50 per cent of the used space leading then to a total of 1.50×200 cycles or 300 cycles. This 300 cycle band is on the basis of 20 cycles per parameter and a single side-band transmitted. This will require equalizing against phase distortion, which will increase the effective delay and will require rather costly phase correcting networks. If these two conditions are not justified for the circuit then there must be transmitted a frequency band wider than a single side-band. At the least this transmitted band might be negligibly wider than a single side-band as in the case of the telephotography filters. At the most it might be doubled for double side-band transmission.

It is well to summarize these estimates of required band width. For high quality speech circuits the minimum band width for speech defining signals is estimated to be as follows:

(1) Where cost is unimportant from 300 cycles at the most down to approximately 100 cycles as a lower limit.

(2) Where cost is an important item from 600 cycles down to 300 cycles.

Thus, it is found that a frequency range of approximately 50 cycles is sufficient for speech defining signals that give all the information as to the instantaneous displacement values or positions of the independent variables of speech production with sufficient accuracy for reproducing reasonably good quality 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 in any narrow frequency range transmission scheme 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 subbands 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 subband. 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 the 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. A specific system involving the case just referred to is described hereinafter.

It is shown in a general schematic form in Fig. 1. Speech coming from the line west reaches the west terminal of the system. For simplicity a four-wire connection obtained with a bridge transformer or hybrid coil and a balancing network N, is shown at this point. The transmission may be over a two-wire circuit, the circuit modifications being obvious. The incoming speech is transmitted through circuit 1 to the west speech frequency compressing circuit 2. Fig. 2 is a detailed showing of this west transmitting circuit or speech frequency compressing circuit 2. The frequency range of the speech is considerably decreased in circuit 2 and then the resulting signals are transmitted over the line L of limited frequency range of transmission to the receiving terminal at the east. Here there is a speech frequency restoring circuit 3 which sets up speech signals more or less the reproduction of the speech signals energizing the west sending circuit 2. Fig. 3 is a detailed showing of this east receiving circuit or speech frequency restoring circuit 3. These speech signals set up by circuit 3 are then transmitted on through speech receiving circuit 4 to the line east. Transmission in the opposite direction is carried on over limited frequency line L' by the same process, the east speech frequency compressing circuit 2' being like circuit 2 except oppositely pointed, and the west speech frequency restoring circuit 3' being like circuit 3 except oppositely pointed.

The simplest circuit to use perhaps is an electrical one where the amounts of power in small frequency subbands of the speech frequency range are used as independent variables. The power in any one such subband is not entirely independent of that in the others but is sufficiently so to insure that we do not need many extra subbands on this account. Such a circuit is relatively easy to build electrically with filters, modulators, equalizers, etc., as circuit elements.

Preparatory to describing Figs. 2 and 3 in detail it is noted that they show a speech frequency range reducing and restoring system having means for reducing the frequency range of speech by analyzing the speech so as to determine or measure its unknown characteristics in terms of speech-defining currents of limited frequency range, a transmitting medium (which may have a limited frequency range of transmission) for transmitting these defining currents or defining signals set up or created by the analyzing means, and means for reconstructing, restoring, simulating, or substantially reproducing the speech in response to the transmitted defining currents or signals.

In order to be a satisfactory electrical definition of the speech signals, the speech-defining currents

or speech-defining signals which are set up to replace the speech signals should not only be sufficiently complete and unique to define the speech signals without error or confusion, but moreover should be practical and convenient, so that a simple automatic analyzing circuit can be used to create or set up the speech-defining signals when a speech wave is applied. Such satisfactory definition is obtained when the speech-defining signals define the speech signals in terms of (1) frequency pattern, and (2) amplitude pattern.

Frequency pattern as used here refers to the number and position of energy-bearing frequency components. There are two distinct types. In one there is a fundamental frequency and all its upper harmonics to several thousand cycles. This fundamental frequency is the rate of vibration of the vocal cords and in general is always increasing (rising inflection) or decreasing (falling inflection) in ordinary speech although it can be held constant by sustaining a sound. In the other type of frequency pattern there is a continuous spectrum of energy so that all frequencies are present, instead of discrete ones in a harmonic ratio, although this second type may be thought of as the limiting case of the first type when the fundamental frequency approaches zero.

Amplitude pattern as used here refers to the distribution of power with frequency from instant to instant, or in other words at any instant. A smooth envelope is taken, as this in connection with the frequency pattern gives a complete specification of the speech sound analyzed.

As will appear from the description of the drawings, about to be given, analyzing speech sounds on the basis of frequency and amplitude patterns is especially convenient because at the receiving or reproducing end of the system the simulating or reproducing devices are separated into two simple, easily recognizable groups. The devices required for producing the frequency patterns will be artificial vocal cords for which a relaxation oscillator is shown and an artificial constriction for which a source of resistance noise is shown. The devices required for producing amplitude patterns will be the artificial lungs controlling the total power put out and the artificial resonances controlling the relative amount in each small frequency band. By working on an absolute basis these characteristics are all handled in a single set of modulating controls shown as gain-controlled amplifiers, the artificial lungs being the power supply batteries or sources and the artificial resonances being filters or tuned networks.

The dual nature of the speech signals as defined electrically leads to a dual type sending or frequency range reducing circuit as shown in Fig. 2 and a dual type receiving or speech frequency restoring circuit as shown in Fig. 3. The speech currents entering the sending circuit 2 energize a frequency pattern control circuit FP and an amplitude pattern control circuit AP. The frequency pattern control circuit comprises but one channel and 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 ten channels AP₁ to AP₁₀ and determines what frequency amplitude pattern we have. The information obtained from these two analyzing elements FP and AP is then expressed in the form of electrical currents which can be used, in modulator Mo in circuit FP and in modulators M₁ to M₁₀ in channels AP₁ to AP₁₀, respectively, to mod-

ulate a number of control current frequencies which are then transmitted through channel band-pass filters C_0 to C_{10} , to the limited frequency line L on essentially a carrier system basis.

At the receiving end of line L these transmitted signals after passing through receiving amplifier RA, are applied to receiving frequency pattern control circuit FP' and receiving amplitude pattern control circuit AP'. In circuit FP', channel band-pass filter C'_0 selects the waves received from channel FP so the selected waves can act, after demodulation in demodulator DM₀, on an energy source of frequency patterns FPS so as to cause currents of the proper frequency pattern to flow from this source.

In circuit AP', channel band-pass filters C'_1 to C'_{10} , in channels AP'₁ to AP'₁₀ respectively, select the currents received from channels AP₁ to AP₁₀ respectively. The selected currents, after demodulation in demodulators DM₁ to DM₁₀ of the respective receiving amplitude pattern control channels AP'₁ to AP'₁₀, are used to control shaping networks SN₁ to SN₁₀ in the respective channels to give the proper amplitude-frequency pattern to the power received from the energy source FPS. We then have our reproduction of the original speech signal for any further transmission in the ordinary manner.

The problem arises of selecting frequency bands to use in the amplitude-frequency characteristic determination. A basis for solution of the problem is found in the function known as the importance function, which has been established by articulation studies. This function varies with the amplitude of speech signal received by the ear. We shall assume that the received speech signal is at the optimum power level. Ten or so independent variables are needed for good speech transmission. Ten such bands for the entire speech frequency range should then be as follows:

Band No.	Importance function	Frequencies	ΔF
1	0 to 0.1	250 to 530 or	280
2	0.1 to .2	530 to 780 or	250
3	.2 to .3	780 to 1100 or	320
4	.3 to .4	1100 to 1500 or	400
5	.4 to .5	1500 to 1950 or	450
6	.5 to .6	1950 to 2350 or	400
7	.6 to .7	2350 to 2900 or	550
8	.7 to .8	2900 to 3750 or	850
9	.8 to .9	3750 to 4950 or	1200
10	.9 to 1.0	4950 to 7100 or	2150

This is not necessarily the best choice of frequencies that can be made. Other considerations enter such as the percentage of increase of frequency in a band, the amount of repetition of pattern in different bands, the noticeable distortion due to smallness or largeness of band, etc. However, these considerations would not greatly affect the final choice of frequencies. Accordingly this set is deemed satisfactory. It is noted that at the most important places the bands are quite narrow in width, 250 cycles, 280 cycles, etc. The three lowest ones are the three smallest and correspond ordinarily to not more than two harmonics each in the case of voiced signals. The next three probably correspond to about three harmonics in the average case of a male voice. The next one which carries us close to 3,000 cycles might correspond to four harmonics. Above 3,000 cycles they widen out quite considerably. On this frequency basis we have altogether then eleven parameters for the entire speech range, one for

the frequency pattern control and ten for the amplitude pattern control. Similarly we require a total of eight parameters to go to 2,900 cycles, nine to 3,800 cycles and ten to 5,000 cycles.

To give a better picture of the proposed circuit it is helpful to indicate frequencies for the various circuit elements. The estimate was made above that for high quality speech transmission the speech defining signals would require transmission frequency bands as follows:

- (1) 100 cycles as the absolute minimum.
- (2) from 300 cycles down to 100 cycles where fairly large costs can be borne.
- (3) from 600 cycles down to 300 cycles where costs must be kept relatively low.

We shall arbitrarily take a 360 cycle total transmission band with other choices such as to tend to typical results after overcoming typical difficulties. This 360 cycle band is on the basis of twelve channels each of 20 cycle pass band with a 10 cycle waste space between. Ten of these are for amplitude control as already chosen, one is for frequency pattern control and one is for synchronizing. We assume single side-band transmission to handle signals including direct current such as is used in telephotography; this means that we must use either linear phase filters of many elements per section or use ordinary filters with complicated phase equalizing networks in addition. (The assumption of a 20 cycle pass band with serious phase difficulties may be a little conservative. However so much of our problem is inherently one of phase difficulties that it is desirable to stress the phase effects.)

On the basis laid down so far we shall proceed to discuss: first, the frequency pattern circuits, second the amplitude pattern circuits and finally the overall circuit.

FREQUENCY PATTERN CONTROL

A particular form of the frequency pattern control circuit is shown in Fig. 2 at the transmitting end and Fig. 3 at the receiving end. Its operation takes advantage of the fact that in vowels and other sounds having a decided fundamental frequency in the range 80 to 320 cycles there is a high power level while in sounds like the sibilant consonants where the power is in a continuous spectrum rather than a discrete one the power level is much lower. When a high level discrete spectrum is applied the frequency pattern control circuit sends to the limited frequency line L a current indicating what the fundamental frequency so applied is, not 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 applied the circuit is not energized so no current from it is transmitted to the line.

Figs. 2 and 3 show the frequency pattern control circuit, with details of the non-linear circuit elements, the other circuit elements being more or less standard.

Referring now especially to Fig. 2, band-pass filter F_0 selects the band from 250 to 500 cycles of the voice signal so as to be sure to include at least two harmonics of speech if the fundamental is below 250 cycles. This is done in case the fundamental may be 80 cycles or so and therefore inefficiently transmitted over the telephone line to the point at which we determine the fundamental frequency. The output of this band-pass filter is fed to detector D which may be merely some small copper-oxide elements. This insures

that a fair amount of the fundamental frequency will be present if the power level is sufficiently high as in the case of vowels. The output from the detector is sent through an attenuating network E_1 , of a type often termed an equalizer, which has a loss increasing with frequency for the purpose of insuring that the fundamental frequency comes out at a higher level than any upper harmonics that may be present. For practical purposes this purifies the fundamental tone. Next the output from this equalizer is fed to a constant output amplifier LA so that from this amplifier there is obtained essentially a single frequency, the fundamental of the speech signal, at a constant power level regardless of what frequency it is.

This fundamental frequency may be from about 80 to 320 cycles. Next we pass this power through an equalizer E_2 similar to the one described previously so that the output from this equalizer decreases as the frequency increases. This output is sent through another copper-oxide detector D_0 . Essentially it gives a direct current bias that fluctuates as the fundamental frequency of speech fluctuates, that is, at syllabic frequencies. The detector output is then sent through a low-pass filter F_{30} cutting off at 20 cycles so that the unwanted higher frequency products are eliminated. This output is now used as a bias on a push-pull amplifier M_0 transmitting a 30 cycle tone obtained from 60 cycle oscillator O, multi-vibrator MV and selective circuit S_0 . This can be considered as a balanced second order modulator, the 30 cycle carrier being modulated by the biasing current as a signal. When there is no energy being received from the low-pass filter F_{30} the bias is such that essentially none of the tone is transmitted to the line.

In proportion as more biasing current is fed from the low-pass filter more of the 30 cycle tone is permitted to pass. The output of this modulator is fed to a band-pass filter C_0 which passes the lower side-band, from 10 to 30 cycles, to the junction point for other control tones, thence to the transmitting amplifier TA and through it to the line.

At the receiving end this control current goes through the receiving amplifier RA shown in Fig. 3, after which it is picked off through a channel band-pass filter C'_0 passing from 10 to 30 cycles. It is then sent through a copper-oxide or other detector DM_0 for the purpose of rectification to syllabic frequencies.

Its rectification is accomplished by demodulating it in DM_0 with 30 cycle current supplied from multi-vibrator MV' through selective circuit S'_0 . The multi-vibrator receives 60 cycle current sent from oscillator O at the transmitting end of the system through filter C_{11} , amplifier TA, line L, amplifier RA, filter C'_{11} and phase shifter PS. This multi-vibrator and the multi-vibrator MV of Fig. 2 have identical circuits generating the harmonics of 30 cycles.

The syllabic frequency currents from demodulator DM_0 are sent through a low-pass filter F'_{30} , 0 to 20 cycles, where the output is used for two purposes. First it biases an amplifier VA which amplifies resistance noise received from resistance R through amplifier A. The bias is so applied to a grid biasing resistor B_1 for the amplifier that when no bias is received the resistance noise from R and A is passed on through VA but as bias is received the output of VA is cut down, that is, the gain of the amplifier VA is decreased by a nega-

tive bias being applied. Secondly, the output of F'_{30} is applied to a biasing resistor B_2 in the common grid lead of a push-pull vacuum tube circuit VR. The purpose here is to control the plate resistances of these tubes. The plate resistances in series are used as the resistance element R_0 of a multi-vibrator circuit MV₀ so that the frequency of the multi-vibrator 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 outputs from the two tubes of the multi-vibrator in series and in parallel and then combine these two so as to generate all the harmonic frequencies. Another possible arrangement of multi-vibrator is one at one-half the frequency from which just the even harmonics are used. When not energized by waves from filter F'_{30} the multi-vibrator circuit stops oscillating. The multi-vibrator output and the resistance noise circuit output from the variable gain amplifier VA are combined in a circuit leading to the amplitude controlling circuits. Preferably the multi-vibrator 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 attained by making the coupling loose between the primary and secondary windings of the multi-vibrator output transformers, and equalizer E_4 omitted.

The attenuation-frequency characteristic of equalizer E_2 determines the variation, with frequency, of the amplitude of the fundamental components (of speech sounds) that are delivered to detector D_0 ; and this input to D_0 has its syllabic frequency component detected by detector D_0 and transmitted by carrier current transmission to the bias resistor B_2 which determines the fundamental frequency of multi-vibrator MV₀. So the voltage output of equalizer E_2 , at each fundamental frequency that E_2 transmits, should be of such value that the voltage across the resistor B_2 will be of proper value to cause the multi-vibrator to assume that fundamental frequency as the fundamental frequency of the multi-vibrator. Therefore, if the voltage transmission-frequency characteristic of the system from the output of equalizer E_2 to the bias resistor B_2 is, for example, flat, or in other words if the ratio of the equalizer output voltage to the resulting voltage across the resistor B_2 is independent of frequency, then the attenuation-frequency characteristic of the equalizer E_2 should be such that the ratio of output voltage to input voltage for the equalizer decreases with frequency increase in the same manner that the fundamental frequency of multi-vibrator MV₀ increases with decrease of the biasing voltage across the biasing resistor B_2 . Then the fundamental frequency set up by MV₀ will increase and decrease in the same manner as the fundamental frequency of speech sound waves in the input of equalizer E_2 .

AMPLITUDE PATTERN CONTROL

The amplitude pattern control circuits are essentially circuits which at the transmitting end 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 end where the output of resistance noise from VA or multivibrator harmonics from equalizer E_4 is shaped accordingly. The fre-

quency bands are chosen as described previously.

The transmitting circuit is shown in Fig. 2, and the receiving circuit in Fig. 3. The ten channels AP_1 to AP_{10} and AP'_1 to AP'_{10} are used to transmit information about the amplitude pattern. Thus the channel 70-90 cycles transmits information about the amplitudes in the speech range 250-330 cycles; the channel 100-120 cycles, in the range 530-780 cycles etc.

Considering the channels AP_1 and AP'_1 , for example, the output from the 250-530 cycle speech band-pass filter F_1 is fed to detector D_1 , which may be, for instance, of the copper-oxide type. The syllabic frequencies in the output from the detector are passed through a 20 cycle low-pass filter F'_{31} and used to bias an amplifier or balanced modulator M_1 transmitting 90 cycles, the amount transmitted depending on the bias. There is no transmission through the amplifier when the bias applied is zero. As the bias builds up it permits transmission. The output from this amplifier or modulator is passed through a 70-90 cycle band-pass filter C_1 to the junction point of the other circuits for amplitude and frequency pattern control, and then through the transmitting amplifier TA, line L, receiving amplifier RA and 70-90 cycle band-pass filter C'_1 to a second order modulator DM_1 that has a 90 cycle carrier supplied from MV' in synchronism with the 90 cycle carrier supplied by MV to M_1 at the sending end. The output of DM_1 is passed through a 20 cycle low-pass filter F'_{31} . The resulting variable direct current is applied to a biasing resistor B_3 to give a grid bias to a signal shaping network or push-pull amplifier SN_1 , which amplifies the signal power received from the multi-vibrator MV_0 or the resistance noise source VA through 250-530 cycle speech band-pass filter F'_1 . This is then fed through a 250-530 cycle speech band-pass filter F''_1 to the input of speech amplifier SA, where the outputs from nine other speech band-pass filters (of channels AP'_2 to AP'_{10}) are combined to give the original speech signal. The speech currents are then transmitted through amplifier SA to the speech receiving circuit 4.

It will be understood that transmitting channels AP_2 to AP_{10} are like AP_1 except as to frequencies involved, and receiving channels AP'_2 to AP'_{10} are like AP'_1 except as to frequencies involved; and that detectors D_0 to D_{10} may be alike, modulators M_0 to M_{10} may be alike, demodulators DM_0 to DM_{10} may be alike, variable gain amplifiers or signal-shaping networks SN_1 to SN_{10} may be alike, filters F_{30} to F_{40} may be alike, and filters F'_{30} to F'_{40} may be alike.

OVERALL CIRCUIT CONSIDERATIONS

While the outputs of detectors D_0 to D_{10} could be transmitted directly over separate lines to filters F'_{30} to F'_{40} , and the illustrated carrier apparatus could then be dispensed with, it is convenient to use the carrier system, especially where the number of frequency bands or channels used for amplitude control is considerable, as for instance in the system shown. In that system the number so used is ten, all employing the common line or transmission medium L between the sending and receiving ends of the system. In a case of this sort it is convenient to supply the required set of carrier frequencies from a single source of carrier waves, such as oscillator O. Further, the same set of carrier frequencies used at the sending end of the system should be transmitted to the receiving end or otherwise ob-

tained synchronously at that end. In the system shown, 60 cycle voltage from oscillator O is sent through multivibrator MV so as to get out a large number of upper harmonics of half of 60 cycles or 30 cycles which are then selected by frequency selective circuits S_0 to S_{10} and used in modulators M_0 to M_{10} as the carrier supply frequencies of the eleven indicated channels FP and AP_1 to AP_{10} . Oscillator O also sends 60 cycle current through filter C_{11} , transmitting amplifier TA, line L, receiving amplifier RA and phase adjuster PS to multi-vibrator MV' , which supplies multiples of 30 cycles from 30 cycles to 360 cycles. These are selected by frequency selective circuits S'_0 to S'_{10} and used as the carrier supply frequencies for demodulators DM_0 to DM_{10} . Preferably the waves of the different frequencies supplied by MV should have equal power levels. The same is true with respect to MV' . The equality can be obtained by amplitude equalization or attenuation equalization in any suitable manner, as for example by making the coupling loose between the primary and secondary windings of the multi-vibrator output transformers, (not shown), as mentioned above for the case of multivibrator MV_0 .

A number of non-linear circuits have been shown. One type, D_0 to D_{10} , is referred to as the detector circuit. In general, its purpose is similar to that of the detectors in voice operated circuits, that is, the received speech power is more or less rectified giving a current component of syllabic frequency to be used as a control. The detector D, for obtaining the fundamental frequency of speech sounds, is similar. Where separate low frequency lines are available the outputs of detectors D_0 to D_{10} can be sent directly to the receiving apparatus. Since, usually, such separate lines are not available, the system shown makes use of different frequency ranges on the same line. Therefore these syllabic frequency control currents are translated to different parts of the frequency spectrum for transmission over the line L of limited frequency transmission range or other common medium of transmission. For this purpose modulation is used so we have a second set of non-linear circuits M_0 to M_{10} to serve primarily as modulators. Similarly at the receiving end we have for the third type, demodulators DM_0 to DM_{10} to give back the control currents. Next the control currents are used for two purposes each of which involves a type of non-linear circuit:

(a) To bias amplifiers VA and SN_1 to SN_{10} and thus change the gain through them;

(b) To vary the plate resistance of vacuum tubes in circuit VR.

Sixthly there are multi-vibrators (relaxation oscillators) for (a) setting up the carrier supply frequencies in MV at the transmitting end of the system and in MV' at the receiving end, and (b) setting up the frequencies in the vowel sounds in MV_0 . Finally there is a constant output level amplifier LA shown in Fig. 2. This makes a total of seven types of non-linear circuit some of which however have much in common.

Delay and delay distortion probably constitute the most serious problem in designing this system. The delay can be equalized at some expense. If however, a wider frequency band is available on the line the delay trouble can be overcome in large degree by transmitting both side-bands. In this case they will fall in the middle of the pass band. The carrier may also

be transmitted or reapplied at the receiving end. Let us consider now easing up in delay distortion will affect the band width required for transmission. The twelve channel circuit discussed here takes a 360 cycle band for high quality speech (7100 cycles) or 270 cycles for 2900 cycle speech which corresponds to commercial telephone circuits. With double side-band transmission this would need to be made about twice as wide, that is, 720 and 540 cycles respectively instead of 360 and 270 as given before. These band widths can of course be adjusted up or down depending on the relative importance of the total frequency band required, of the total delays permitted, the quality desired, the permissible cost etc. A number of things can be done to reduce the delay distortion such as using less frequency space for the useful band and more for the waste space between channels, having less selectivity, etc.

In the circuit design it is important to see that the delay in all the different branches is the same. Since the frequency pattern control branch tends to have more inherent delay than the amplitude control branches it is desirable to have a certain amount of delay in common with all the amplitude control circuits as is indicated by delay equalizer DE in Fig. 2 and delay equalizer DE' in Fig. 3.

Two sets of band-pass filters F_0 to F_{10} and C_0 to C_{10} , are shown at the transmitting end and two, C'_0 to C'_{10} and F''_1 to F''_{10} at the receiving end in addition to the selective carrier supply circuits and the low-pass filters. If desired, a third set, F'_1 to F'_{10} may be used at the receiving end, making twenty, instead of ten speech band filters in the receiving amplitude control circuits, for the case of the 7100 cycle band to be transmitted. These ten or twenty speech band filters might have roughly requirements of about 6 decibel loss discrimination at the separating points between two bands, 20 decibel loss at the middle of the next band and 40 decibel loss at the middle of the second band. Assuming the requirements to be approximately these, the delay in the speech band filters is relatively small, two of them in tandem giving only about ten milliseconds delay. These speech band filters can probably have much easier attenuation requirements than indicated here. Those at the receiving end in particular might well be made to be more nearly tuned circuits than sharp filters. This will tend to put out a rounded amplitude-frequency envelope for speech at the receiving end rather than one with sharp corners.

The set of channel filters at each end, that is, filters C_0 to C_{10} and C'_0 to C'_{10} , are spaced much closer together in the frequency range and therefore inherently tend toward much larger delays. Assuming they require about 6 decibels loss discrimination at the edges of the pass bands, 20 decibels loss at the beginning of the next band and 40 decibels loss at the middle of the next band we get a set of filters that are estimated roughly to have about 100 milliseconds delay at the edges and 50 in the middle. With this it would be necessary to equalize this delay if single side-band is used. If this is done the total delay in the band will probably be increased to about 150 milliseconds. The resulting filter will also be fairly costly as delay correcting networks require many elements. It is quite probable that these channel filters can be used with much easier requirements also. If this system is used

where wider frequency bands on the transmission lines are available, it will ordinarily be desirable to use double side-band transmission. There are other elements in the circuit that cause a certain amount of delay and therefore delay distortion such as the low-pass filters F_{30} to F_{40} and F'_{30} to F'_{40} and the constant output amplifier LA. These, however, will total up only to rather small values. The total overall one-way delay on the basis of 20 cycle channel bands with 10 cycle separation will amount to not over .35 second.

A circuit of this sort at the receiving end needs good equalization of both attenuation and delay. The adjustment of the amplitude control circuits can be made by transmitting a single tone in the channel under test.

If desired, to simplify delay equalization in the system, the speech band-pass filters may be made of constant band width, as for example 300 cycles. For instance, for transmitting 250 to 2950 cycle speech, nine amplitude pattern control channels have been used, with filters F_1 to F_9 respectively passing the speech bands 250-550 cycles, 550-850 cycles, 850-1150 cycles, 1150-1450 cycles, 1450-1750 cycles, 1750-2050 cycles, 2050-2350 cycles, 2350-2650 cycles, and 2650-2950 cycles. Then filters F'_1 to F'_9 respectively passed these same bands, and likewise filters F''_1 to F''_9 respectively passed these same bands.

Another simplification of the circuit of some interest is obtained by dropping out the frequency pattern control channel FP and FP'. We then cannot switch from the multi-vibrator circuit MV₀ to the resistance noise source FPS, but must use either the resistance noise or the multi-vibrator now set at an arbitrarily chosen frequency such as the average vocal cord frequency. Interesting sound effects are thus produced. With the resistance noise source the manufactured speech has the sound of a whisper. With the fixed frequency multi-vibrator source the manufactured speech has the monotone of a chant. Both circuits, however, give good intelligibility at a considerable reduction in the circuit equipment required and may therefore be very useful where naturalness is not an important factor or not a requirement in transmitting speech sounds.

Ordinarily, with a system of the type shown the permissible range of volume transmitted might well be somewhere from 20 to 40 decibels for example. The wider the frequency band the easier it is to get a large number of volume range steps at the receiving end without too fine adjustments of the apparatus. The smaller the volume range, the more readily can transmission be carried on over long circuits of high attenuation such as submarine cables. That is, signal levels down much nearer to resistance noise can be used and therefore the cable can be a cheaper cable of greater attenuation than could otherwise be used, or a wider frequency range can be used on a given cable.

Any desired portion of the circuit or link between the terminals of the system may be a two-wire, four-wire or radio circuit. It will be understood that for multiplex carrier transmission of a plurality of conversations simultaneously over a common transmission circuit or medium, either by wire or by radio transmission, each speech channel may have its speech band reduced in frequency range by a speech frequency compressing circuit (such as the circuit 2 of Figs. 1 and 2) and then applied to the modulating channel of

the wire carrier or radio carrier multiplexing apparatus at one end of the system. At the other end the band detected by the demodulator of the wire carrier or radio carrier multiplexing apparatus will correspond to this narrowed band and may be impressed upon a speech frequency restoring circuit (such as the circuit 3 of Figs. 1 and 3).

It is emphasized that the output of the speech frequency compressing circuit 2 of Figs. 1 and 2 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 medium L, directly contain no intelligibility whatever.

Where the circuit connecting the terminals of the transmission system is, for example, a trans-Atlantic submarine telephone cable, the reduction of the frequency range transmitted can be of great aid especially because of the difficulty of inserting the desired number of amplifiers to keep down the attenuation at high frequencies.

In multiplex telephony over cable or open wire lines or by radio transmission, the reduction of the frequency range can facilitate an increase of the number of channels on a line or in the transmission medium, as for example the placing of two or more speech channels, on the line or in the transmission medium, in the frequency range ordinarily occupied by one.

The compression of the frequency range of the transmission in accordance with the invention, can increase the signal-to-noise ratio obtaining at the receiver without necessitating increase of the sending level of the signal. This is of especial importance in systems, for example, long radio transmission systems such as the trans-Atlantic, in which the transmission medium is unsatisfactory for the desired grade of transmission of the original communication signals because of inadequate signal-to-noise ratio in the transmission medium.

It is seen that the system described above, instantaneously analyzing speech waves, and then, upon the basis of the information obtained by the analysis, synthesizing or reconstructing the speech, illustrates doing electrically what one does in mimicking when he listens to speech sounds with his ear and immediately reproduces them with his own vocal system. In this process three operations may be distinguished: first, the aural perception and analysis of the sound; second, the passage of neural impulses for the control of the vocal organs; and third, the reproduction of the sound. In the system described above these operations are electrical, or may be considered electro-acoustical if the microphone applying the electrical speech waves to be analyzed and the loud-speaker or telephone receiver actuated by the reconstructed electrical sound wave are regarded as included in the system.

These operations are to be distinguished from those of ordinary telephony. In telephony there is neither analysis nor synthesis of the sound. Instead, a listening device, the telephone transmitter, picks up speech sounds and translates them into current variations; a complex alternating current, with components corresponding to those of the sound wave, is then transmitted to the other terminal; and there a reproducing, or talking, device in the form of a telephone receiver reproduces the speech. In ordinary telephony a sound wave is electrically moved from one point to another by direct transmission; but

when the synthesizing process described above is to be used, only the specifications for reconstructing the sound wave are directly transmitted.

By the analyzing device in the system described above, the electrical speech wave coming from the microphone is instantaneously analyzed for the three characteristics, pitch, resonant frequency regions and loudness; control currents are then passed to the speech synthesizer; in this a new current is produced, manufactured according to the specification of the speech analyzer; and this current is suitable for actuating a telephone receiver to produce acoustical speech waves.

Thus the system between the microphone and the telephone receiver analyzes a sound-bearing electrical current and synthesizes a mimicking electrical current. On this basis the analyzing device is the artificial ear of the mimicking system and the synthesizing device is the artificial vocal system. On the other hand, if one thinks of an artificial ear as a device for analyzing an acoustic speech wave, then the microphone is to be regarded as part of the artificial ear. Likewise from the acoustical standpoint the artificial vocal system would be regarded as including the loud-speaker or telephone receiver.

Of the two basic elements, the analyzer and the synthesizer, the analyzer is like the ear in that it analyzes speech sounds to determine their pitch, resonant frequency regions and loudness.

In the synthesizer described in detail above, the element equivalent to the vocal system is entirely electrical, and produces an electrical current. This synthesizer has electrical equivalents of all of the important elements of the vocal system. To bring this out, the operation of the vocal system should be considered. The three fundamental steps in producing speech sounds are: (1) providing an energy source; (2) modulating energy from this source so that it will contain vibrations in the frequency range audible to the ear; and (3) selecting from these audible vibrations by resonance to get the various speech sounds.

The energy source is provided by the pressure on the air in the lungs as it is being exhaled. In a normal exhalation this flow of air is not audible. It is made audible by forming a partial closure and forcing the air through under pressure. The air particles which get through then vibrate in rapid complex motions at frequencies that are audible to the ear. For voiced sounds, this partial closure is at the vocal cords which then vibrate at their natural periodic rate, depending on the tension the talker places on them. For unvoiced sounds the closure is in the mouth passage. Thus "s" is formed by placing the tongue against the hard palate and forcing the air through the narrow opening between them. After these audible vibrations are set up it is still necessary to provide means for differentiating the various sounds from one another. This is done by shaping the mouth to form resonant air chambers which favor certain frequencies but discriminate against others. It is in this way that "ah" differs from "ee" although in both cases the vocal cords are vibrating.

Starting from the lungs and going out toward the mouth there are (1) the steady power source in the lungs, (2) the vibration producers in the form of the vocal cords for the voiced sounds and the constrictions in the mouth passage for the unvoiced sounds, and (3) the resonance elements in the form of resonant air chambers.

The artificial speech synthesizer shown in the drawings has elements corresponding to these

three elements of the vocal system as follows: (1) The steady power supply may be thought of as from batteries instead of compressed air in the lungs. (2) The vibrating elements are the multi-vibrator for the vocal cords and the resistance noise source for the unvoiced sounds produced at the constriction in the mouth. (3) Resonance or frequency discrimination, which determines the frequency components to be favored and those to be discriminated against, is provided by tuned or frequency selective electrical networks instead of air chambers as in the mouth. The output is an electrical wave, instead of an acoustic wave.

When a speech signal is analyzed by the artificial ear it is resolved into simple speech-defining signals or control currents. However, as indicated above, these currents can be modified or controlled by controlling the circuit elements in their paths and thus the manufactured speech can be made to differ in prescribed manner from the original speech.

The invention is capable of embodiment in widely different systems.

What is claimed is:

1. In communicating intelligence of a speech signal having predictable characteristics known to be in the signal and unpredictable variable characteristics, neither said predictable nor said unpredictable characteristics being adequate to give intelligibility to the signal, the method which comprises transmitting waves of frequency below audibility representing the unpredictable characteristics through a transmitting medium, receiving those waves from the medium to the exclusion of waves representing the predictable characteristics, generating waves not received from the transmitting medium that define the predictable characteristics, and so combining effects of the transmitted waves received from the transmission medium and the generated waves not received from the transmitting medium as to create a synthetic speech signal containing substantially the same intelligence as the original speech signal.

2. A signaling system comprising means supplying signals that contain modulated waves whose modulations correspond to other waves of lower frequency, the unknown characteristics of said signals being definable by said other waves and a set of unmodulated waves, neither said other waves nor said unmodulated waves being sufficient to give intelligibility of the signals, means deriving from said signals said other waves to the exclusion of said unmodulated waves, wave generating means supplying under control of said derived waves a set of waves of frequencies not received by said generating means from said signals but comprised in the frequency range of said unmodulated waves, and means responsive to said derived waves and said generated waves for reproducing the signals.

3. A signaling system comprising means deriving, from speech signal waves having predictable characteristics and unpredictable characteristics, waves of frequency below the fundamental frequency of vibration of the vocal cords defining the unpredictable characteristics to the exclusion of the predictable characteristics, means for generating waves not received thereby defining said predictable characteristics, and means for translating said derived waves and said generated waves into waves representing the speech signals.

4. The method which comprises analyzing

communication signal waves from the standpoint of oscillatory and modulatory characteristics, transmitting to a receiving point waves representing the modulatory characteristics and having a smaller product of frequency band width and time than said signal waves, producing locally at the receiving point waves corresponding to the oscillatory characteristics and having frequencies not present in the waves transmitted to the receiving point, and utilizing the produced waves and the transmitted waves to create waves serving as a substitute signal for the signal represented by the waves analyzed.

5. A communication system having a sending end and a receiving end for communicating a signal which is produced by a signal source and whose production involves mechanical motions, said system comprising means at the sending end of the system for analyzing the signal into variable features and features of relatively fixed nature and further analyzing the variable features into simpler signals for defining the motions involved in producing the signal to be communicated, said defining signals having the summation of their products of frequency band width by time smaller than the product of frequency band width by time of said signal to be communicated, means for transmitting from the sending end of the system to the receiving end said defining signals to the exclusion of information of said features of fixed nature, and at the receiving end of the system a wave source supplying said features of fixed nature independently of the first-mentioned source, and means controlled by said transmitted defining signals for reproducing said variable features.

6. A system for communicating a signal whose production involves mechanical motions comprising means for analyzing the signal into features of variable nature and features of relatively fixed nature and further analyzing the features of variable nature into signals having a smaller product of frequency range and time for defining the motions involved in producing the signal to the communicated, means for reproducing the features of relatively fixed nature independently of the signal analyzed, and means controlled by said reproducing means and said defining signals for reproducing the signal to be communicated.

7. The method of treating a speech signal whose generation involves sluggish muscular movements causing variations in acoustic pulsations resulting from rapid vibration of vocal cords, which comprises deriving, from said signal, waves of frequency lower than said acoustic pulsations, that define said variations but not said pulsations.

8. The method which comprises dividing a frequency band of waves into subbands respectively having syllabic time rates of change of energy, measuring said respective rates, generating waves having a discrete energy spectrum with substantially the same power at each frequency of energy concentration, and varying at said respective measured rates the energy in respective frequency bands of said generated waves.

9. 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 discrete energy spectrum and having the relative amplitudes of their components independent of the relative amplitudes of the components of said transmitted energy, and means controlled by the energy in said channels for varying at said rates, respec-

tively, the energy in different frequency bands of waves from said source.

10. A system for communicating a message represented by a frequency band of waves, comprising a source of waves having a discrete energy spectrum and having the relative amplitudes of their components independent of the relative amplitudes of the components of the first-mentioned waves, analyzing circuits responsive to said frequency band of waves for determining the syllabic time rate of change of energy in each of a plurality of subbands of said frequency band of waves, and means controlled by said analyzing circuit for producing in corresponding frequency subbands of the output from said source a similitude of these time rates of change of energy.

11. The method which comprises dividing a frequency band of waves that represents a voice wave into subbands respectively having syllabic time rates of change of energy, measuring said respective rates, producing a complex tone having its fundamental frequency equal 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, and varying at said respective measured rates the energy in respective frequency bands of said complex tone.

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 equal 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 syllabic time rate of change of energy in each of said subbands, and means controlled by said analyzing circuits for producing from the energy of said complex tone a similitude of said time rates of change of energy in said subbands.

13. The method which comprises dividing a frequency band of waves into subbands respectively having syllabic time rates of change of energy, measuring said respective rates, generating independently of said waves other waves having a continuous frequency spectrum and substantially uniform distribution of power with frequency over the frequency range of said bands, and varying at said respective measured rates the energy in respective frequency bands of said generated waves.

14. A system comprising a plurality of wave transmission channels for respectively transmitting subbands of a frequency band of waves that represents a voice wave, a source of waves representing resistance noise, analyzing circuits responsive to waves from said channels for determining the syllabic time rate of change of energy in each of said subbands, and means controlled by said analyzing circuits for producing in the output from said source a similitude of said time rates of change of energy in said subbands.

15. The method of expressing the information content of speech signals in a compressed frequency range which comprises setting up an energy pattern and varying it in a limited number of respects, not exceeding approximately ten, corresponding to a limited number of muscular changes used in speech production and with a limited time rate of change in each respect cor-

responding to the limited rate, less than the fundamental rate of vibration of vocal cords, at which the speech producing muscles can be varied in producing speech.

16. The method which comprises creating an energy pattern varying in a limited number of respects corresponding to a limited number of muscular changes used in speech production and with a limited time rate of change in each respect corresponding to the limited rate, less than the fundamental rate of vibration of vocal cords, at which the speech producing muscles can be varied in producing speech, generating waves having a discrete energy spectrum with substantially the same power at each frequency of energy concentration, and varying the amounts of power at the different frequencies in said waves in accordance with variations in said energy pattern.

17. The method which comprises deriving from a voice wave an energy pattern varying in a limited number of respects of the order of ten corresponding to a limited number of muscular changes used in speech production and with a limited time rate of change in each respect corresponding to the limited rate, not exceeding approximately ten cycles per seconds, at which the speech producing muscles can be varied in producing speech, deriving from said voice wave a wave having a frequency equal to the fundamental frequency of the vibrations of the vocal cords of the talker and waves of amplitudes equal to the amplitude of the last-mentioned wave and frequencies of upper harmonics of the last-mentioned wave, and varying the amounts of power at the different frequencies in said derived waves in accordance with variations in said energy pattern.

18. A system comprising means for creating an energy pattern varying in a limited number of respects of the order of ten corresponding to a limited number of muscular changes used in speech production and with a limited time rate of change in each respect corresponding to the limited rate, not exceeding approximately ten cycles per second, at which the speech producing muscles can be varied in producing speech, a source of waves generating independently of the energy of said energy pattern waves having a continuous energy spectrum and substantially uniform distribution of power with frequency, and means responsive to the variations in said energy pattern for varying the amounts of power at the different frequencies in said waves in accordance with variations in said energy pattern.

19. The method of obtaining secrecy in communication which comprises transforming a set of communication signals containing modulated waves whose modulations correspond to other waves of lower frequency into a set of defining signals defining said other waves but not said first-mentioned waves, and transmitting to a distance said defining signals to the exclusion of both said communication signals and said first-mentioned waves as a substitute for said communication signals.

20. A privacy system for speech transmission comprising means for translating speech signals containing a complex modulated oscillatory wave whose modulations correspond to a complex modulatory wave into a set of signals defining said modulatory wave but not said oscillatory wave and exclusive of said modulatory wave and directly containing no intelligibility whatever, a transmission medium for transmitting said de-

fining signals to the exclusion of said speech signals and said oscillatory wave, and means responsive to said transmitted defining signals for reproducing said speech signals, said defining signals having the summation of their products of frequency band widths by time less than the product of frequency band by time of the reproduced speech signals.

21. A system for communicating a speech signal comprising means for transforming said signal into waves defining the sluggish muscular movements involved in the production of the speech signal, means for transmitting said waves to the exclusion of waves corresponding to acoustic pulsations resulting from the rapid vibrations of vocal cords in producing said signal, and means responsive to said transmitted waves for reconstructing said signal.

22. The method which comprises analyzing a complex communication signal wave having known oscillatory characteristics and unknown variable modulatory characteristics into defining signals exclusive of every frequency component of the complex signal wave which do not define the known oscillatory characteristics of the complex signal wave but define the unknown variable modulatory characteristics of the complex signal wave required to reconstruct the signal.

23. A speech transmission system comprising means responsive to speech signals for producing waves whose frequency components differ from each of those of said speech signals in form and in position in the spectrum and have less total frequency range than that of said speech signals and certain of which define the amplitude pattern and certain others of which define the type of the frequency pattern of said speech signals, and means responsive to said produced waves for reproducing said speech signals.

24. The combination of a system for transmitting a complex signal and means for supplying said signal to said system, said system comprising means for subdividing said supplied signal into frequency subbands of the signal frequency band, a source of waves of frequency subbands corresponding to said frequency subbands of the signal, respectively, and means responsive to the waves of the latter frequency subbands for varying the power of the frequency subbands of the waves from said source at the syllabic time rates of change of the power in the respectively corresponding frequency subbands of said signal.

25. In combination, a source of speech signal waves and means connected to receive said waves and producing in response thereto waves having amplitude variations corresponding to frequency variations of the fundamental frequency component of said speech signal waves.

26. In combination, a source of waves having amplitude variations corresponding to frequency variations of the fundamental frequency component of speech sounds, and means for translating said amplitude variations into frequency variations.

27. In combination, a source of waves having amplitude variations corresponding to frequency variations of the fundamental frequency component of speech sounds, and means responsive to said waves for generating waves of the frequencies of said component and harmonics of the last-mentioned waves.

28. The method of reproducing speech signals which comprises deriving therefrom defining waves which define the unpredictable characteristics of the speech signals and whose product of

time duration and total frequency range has an order of magnitude inferior to that of the speech signals whose unpredictable characteristics are defined by said waves, and reproducing said signals from said waves.

29. In combination, means for producing a complex wave having a fundamental component of varying frequency and harmonics thereof each of substantially the same power as said fundamental component, means for separately amplifying waves of different frequency subbands of said complex wave, means for varying the amount of the amplification of the waves of the respective subbands at frequencies low compared to the frequency of said fundamental component, and means for combining the amplified waves to form a resultant wave.

30. In combination, means for producing a complex wave having a continuous energy spectrum and having substantially uniform distribution of power with frequency, means for separately amplifying waves of different frequency subbands of said complex wave, means for varying the amount of the amplification of the waves of the respective subbands, and means for combining the amplified waves to form a resultant wave.

31. In combination, means for selecting a subband of a speech band of waves in the lower portion of the band, a circuit responsive to said subband for obtaining the fundamental frequency component of the band to the exclusion of other components, said circuit comprising a detector responsive to said subband and, in tandem therewith, equalizing means having attenuation increasing with frequency and a constant output level amplifier, a detector responsive to waves from said circuit, and a low-pass filter fed from said second detector.

32. A two-way telephone transmission system comprising a transmitting medium unsatisfactory for transmission of the original speech signals because of their frequency band width, and two two-way terminals each for transmission to and reception from said medium; each of said terminals comprising a two-way speech frequency transmission circuit and a transmitting circuit and a receiving circuit connected between said speech frequency transmission circuit and said medium; each of said transmitting circuits comprising a speech frequency band compressor and each of said receiving circuits comprising a speech frequency band restorer; each of said compressors comprising a transmitting frequency pattern control circuit and a transmitting amplitude pattern control circuit, and each of said restorers comprising an energy source of frequency patterns, a receiving frequency pattern control circuit and a receiving amplitude pattern control circuit; at one of said terminals the transmitting frequency pattern control circuit comprising means connected to the speech frequency transmitting circuit for selecting from a speech frequency band of waves a frequency subband from which the fundamental component of the speech signals can be derived, a circuit for deriving said fundamental component to the exclusion of other waves, means for detecting the syllabic frequency amplitude variations of said fundamental component, and means responsive to said variations for transmitting to said medium waves representing said variations; at the other of said terminals the receiving frequency pattern control circuit comprising means responsive to the latter waves for reproducing said variations and the energy source of

frequency patterns comprising a source of power having a continuous spectrum, a source of power adding its output energy to that of said continuous spectrum source and having a discrete spectrum including a variable frequency fundamental component and harmonics thereof, and means responsive to said reproduced amplitude variations for rendering said continuous spectrum source ineffective, rendering said discrete spectrum source effective, and controlling the frequency of said fundamental and harmonic components in accordance with the frequency variations of said derived fundamental component of the speech signals at said one terminal; at said one terminal the transmitting amplitude pattern control circuit comprising a number of transmitting amplitude pattern control channels of the order of ten and a delay equalizer connecting them to said speech frequency transmission circuit; each channel comprising a speech frequency subband selector for selecting a different subband of said speech frequency band of waves, means for detecting syllabic frequency variations of the power in such subband, means for translating the detected waves to a given position in the frequency spectrum, said positions being different for the respective channels, and a channel frequency band selector for transmitting the translated waves to said medium; at said other terminal the receiving amplitude pattern control circuit comprising a delay equalizer and receiving amplitude pattern control channels corresponding respectively to said transmitting amplitude pattern control channels, for receiving waves transmitted thereto from said transmitting amplitude pattern control channels through said medium; each receiving amplitude pattern control channel comprising a receiving amplitude pattern control channel frequency band selector selecting the band transmitted by the channel frequency band selector of the corresponding transmitting amplitude pattern control channel, means for restoring the waves received from the receiving amplitude pattern control channel frequency band selector to their original position in the frequency spectrum, means responsive to the restored waves for controlling the power level in a subband of the frequency band of waves from said energy source of frequency patterns and transmitting the controlled subband of waves to the speech frequency transmission circuit, said subband having the frequency limits of the speech frequency subband selector of the transmitting amplitude pattern control channel corresponding to the receiving amplitude pattern control channel.

33. The method of reproducing a signal containing variable information and invariable information and represented by a complex wave which comprises deriving from the complex wave a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the signal source, deriving from said parameters a set of defining waves that respectively define the variations of said parameters, transmitting through a wave transmission medium said defining waves to the exclusion of waves representing said invariable information, producing artificially waves representing said invariable information to the exclusion of said variable information, and combining effects of said artificially produced waves and said transmitted waves to reproduce said signal.

34. The method which comprises analyzing a speech signal into two factors one of which represents the natural flow of energy from the mechanical-acoustical oscillating system comprising the vocal system of a man and the other of which represents the rational control applied to vary this flow of energy so as to impress an information content on it, transmitting signals derived essentially from the second factor alone to a distant point, and there originating an energy stream equivalent to said natural flow of energy before said information content was impressed on it and using the transmitted signals to control the energy stream in such way as to reproduce the original speech signals.

35. The method of analyzing a signal having a fundamental component of varying frequency and harmonics thereof lying in the frequency range of the varying fundamental component, which comprises transmitting the harmonic and fundamental components of a frequency range including a substantial frequency range of the harmonics with the efficiency of said transmission decreasing as frequency increases throughout the latter frequency range, and deriving from the transmitted waves pulses of uniform amplitude proportional in number to the periods of said fundamental component.

36. The method of analyzing and reconstructing a speech signal which comprises selecting the varying frequency fundamental component of the voiced speech sounds to the exclusion of the other components of the speech sounds and utilizing said selected component to create a wave consisting essentially of a fundamental component of said varying frequency and harmonics thereof.

37. The method which comprises producing waves containing the fundamental frequency component of waves representing voiced sounds by intermodulation of upper harmonics of the varying frequency fundamental component, and suppressing the harmonics by transmitting the modulation products with transmission efficiency decreasing as frequency increases over a substantial frequency range including a major portion of the frequency range common to the varying frequency fundamental component and its harmonics.

38. The method of reproducing signals which comprises deriving, from the particular signal being reproduced, waves which are a function of said signal, producing under control of said signal, waves of character dependent upon the character of the signal source but not upon the particular signal being reproduced, and combining the first-mentioned waves and the second-mentioned waves to reproduce the signal.

39. The method which comprises creating a set of waves defining the frequency-time content of a signal, creating a set of waves defining the amplitude-time pattern of the signal in a series of frequency bands, the summation of the products of the frequency bands by time of said two sets of waves being less than the product of the frequency band by time of said signal, and utilizing said two sets of waves to construct said signal.

40. In combination, a source of waves defining the frequency-time content of a signal, a source of waves defining the amplitude-time pattern of the signal in a series of frequency bands, the summation of the products of the frequency bands by time of said two sources being less than the product of the frequency band by time of said signal, and means controlled by said two sources for reconstructing said signal.

41. A privacy method for speech transmission

which comprises deriving from speech unintelligible signals defining syllabic frequency variations of energy in a set of adjacent frequency subbands of speech, transmitting these defining signals, and reconstructing speech by using the defining signals to control local sources of energy.

42. A privacy system for speech transmission comprising means deriving speech defining signals in terms of time flow of energy in a set of adjacent frequency subbands of speech, means for transmitting these unintelligible defining signals, a harmonic producer and a generator of a continuous spectrum of resistance noise, and means for reconstructing speech by using said transmitted defining signals to maintain the fundamental frequency of said harmonic producer in tune with the speech fundamental and to control alternately in accordance with said derived signals the time rate of energy output of said harmonic producer or of said generator.

43. A transmission system comprising a source supplying a message signal which may alternately consist of frequencies in an organized relationship and frequencies in a disorganized relationship, means for generating and transmitting information signals corresponding to the time energy flow in a series of bands with the message signal in either form, two corresponding local sources of energy at the receiving end, one of organized frequency content and the other of disorganized frequency content, and means for reconstructing the message signal at the receiving end responsive to said transmitted information signals for automatically controlling the energy from either of said two local sources.

44. In a signaling system, means to produce a signal containing variable information and invariable information represented by a complex wave, means to analyze said wave and derive therefrom a simple set of parameters having approximately the number of degrees of freedom of the variable elements of said signaling source, means to translate said set of parameters into a set of defining waves that respectively define the variations of said parameters, means for transmitting said defining waves over a transmission medium, means for excluding from the transmission medium waves representing said invariable information, means for producing artificially waves representing said invariable information to the exclusion of said variable information, and means for combining effects of said artificially produced waves and said transmitted waves to reproduce said signal.

45. The method of reproducing a signal containing variable information and invariable information and represented by a complex wave which comprises deriving from the complex wave a simple set of parameters having approximately the number of degrees of freedom of the variable elements of the signal source, deriving from said parameters a set of defining waves that respectively define the variations of said parameters, the frequency of said defining waves being below audibility, transmitting through a wave transmission medium said defining waves to the exclusion of waves representing said invariable information, producing artificially waves representing said invariable information to the exclusion of said variable information, and combining effects of said artificially produced waves and said transmitted waves to reproduce said signal.

46. In a signaling system, means to produce a signal containing variable information and invariable information represented by a complex

wave, means to analyze said wave and derive therefrom a simple set of parameters having approximately the number of degrees of freedom of the variable elements of said signal source, means to translate said set of parameters into a set of defining waves that respectively define the variations of said parameters, the frequency of said defining waves being below audibility, means for transmitting said defining waves over a transmission medium, means for excluding from the transmission medium waves representing said invariable information, whereby the frequency range employed for transmission over said medium is only a fractional part of the band width required for the original signal, means for producing artificially waves representing said invariable information to the exclusion of said variable information, and means for combining effects of said artificially produced waves and said transmitted waves to reproduce said signal.

47. The method of operating upon speech which comprises reproducing the speech and varying the relation of the fundamental frequency of the reproduced speech to the rapidity of the reproduced speech from the relation existing between the fundamental frequency and the rapidity of the original speech.

48. The method which comprises artificially producing speech and fixing the fundamental frequency of the produced speech.

49. The method of controlling the character of synthesized speech which comprises increasing the fundamental period of speech relatively to the duration of the speech.

50. The method of producing a prescribed change in the character of speech that has a varying fundamental frequency, which comprises reproducing the speech and fixing the fundamental frequency of the reproduced speech at a frequency value approaching zero as a limit while preserving in the reproduced speech the intelligibility of the original speech.

51. A wave translating system comprising means supplying a wave having a fundamental component of varying frequency and harmonics thereof lying in the frequency range of the varying fundamental component, means for transmitting said wave having transmission efficiency decreasing with frequency increase throughout a substantial frequency range of said harmonics, means responsive to the transmitted wave for translating it into a wave comprising pulses proportional in number to the periods of said fundamental component but substantially equal to each other in amplitude, and means responsive to said pulses for producing a current whose magnitude is a prescribed function of the varying frequency of said fundamental component.

52. The method of producing a signal which comprises generating one set of waves and modulating it with another set, said one set occupying substantially the same portion of the frequency spectrum as said signal, and said other set having its top frequency substantially lower than the top frequency of said signal.

53. The method which comprises generating one set of waves and modulating each of its frequency components with some portion of the frequency components of another set to produce a signal, said one set occupying substantially the portion of the frequency spectrum occupied by the signal, and said other set having its products of frequency band widths and times sum to less than those of the signal.

54. A signaling system comprising means supplying signals that are modulated waves definable by two sets of waves, neither set being sufficient to give intelligibility of the signals, means deriving from said signals one of said sets to the exclusion of said other set, and means for generating said other set and modulating it with said one set to produce a simulation of the signals, said one set having its top frequency substantially lower than the top frequency of the produced simulation of the signals.

55. The method which comprises supplying signals definable by two sets of waves, one set defining frequency and amplitude modulation of the other set, neither said one set nor said other set being sufficient to give intelligibility of the signals, deriving from the signals a third set of waves having variations definable by said one set, detecting said variations to obtain said one set to the exclusion of said other set and generating said other set and modulating it with said one set to produce a simulation of the signals.

56. The method which comprises supplying signals having a fundamental frequency and harmonics thereof both modulated in frequency and amplitude, said signals being definable by two sets of waves, one set defining frequency and amplitude modulation of the other set, neither said one set nor said other set being sufficient to give intelligibility of the signals, deriving from said signals said one set to the exclusion of said other set, and generating said other set and modulating it with said one set to produce a simulation of the signals.

57. The method which comprises analyzing communication signal waves into two sets of waves, one set defining modulation of the other set, said one set having its products of frequency band widths and times sum to less than the product of frequency band widths and times of said signal waves, transmitting said one set to a receiving point, and there generating said other set and modulating it with said one set to produce a simulation of the signal.

58. The method which comprises instantaneously resolving speech into two sets of waves, the sum of the widths of all of the frequency bands of the two sets being less than the frequency range of the speech*resolved, one set defining the pitch of the speech and the other its distribution of power with respect to frequency.

59. The method which comprises instantaneously deriving from electrical speech waves by selection and demodulation of components thereof two sets of waves, one set defining the pitch of the speech and the other defining its frequency distribution of power, the sum of the widths of all of the frequency bands of the two sets being less than the frequency range of the speech waves, and generating a third set of waves and modulating them with said two sets of waves to synthesize waves simulating the speech waves.

60. The method which comprises instantaneously deriving from electrical speech waves by selection and demodulation of components thereof two sets of waves, one set defining the pitch of the speech and the other defining its frequency distribution of power, the sum of the widths of all of the frequency bands of the two sets being less than the frequency range of the speech waves, and synthesizing speech-simulating waves by generating a new set of waves and utilizing said two sets of waves respectively for

frequency modulation and amplitude modulation of components of said new set.

61. The method which comprises selecting frequency subbands of a band of components of speech and demodulating the subbands to obtain a set of low frequency waves whose products of frequency range and time sum to less than the product of frequency range and time of the band of speech components, said waves corresponding respectively to said subbands, and synthesizing speech by generating a new set of waves having subbands of components that respectively have the same frequency ranges as the first-mentioned subbands and modulating the subbands of the new set of waves respectively with the low frequency waves that correspond to the first-mentioned subbands of like frequency ranges.

62. The method of artificially creating speech from unintelligible audible frequency waves which comprises artificially producing relative variations in the time flow of energy in frequency subbands of the waves substantially in accordance with the relative variations of the time flow of energy that are respectively characteristic of corresponding subbands of the frequency components of speech sounds to be created.

63. The method of artificially creating speech from unintelligible audio frequency waves which comprises separating frequency subbands of the waves and producing by individual amplitude control of the separate subbands of waves relative variations in the time flow of energy in the subbands substantially in accordance with the relative variations of the time flow of energy that are respectively characteristic of corresponding subbands of the frequency components of speech sounds to be created.

64. In a system for artificially producing speech, a source of sound frequency vibrations, means operating at the syllable rate of uttered speech for producing modulations in the total output from said source, a sound reproducer, and a plurality of separate selective devices for selectively controlling transmission of selective frequency ranges of said modulated output from said source to said reproducer to simulate speech.

65. In a system for artificially producing speech, a source of fundamental and harmonic frequency vibrations extending over the essential speech spectrum, a plurality of selective devices for dividing the frequency range into comparatively narrow bands each including certain of said vibrations, an individual loudness control associated with each of said selective devices for controlling the amplitude of the vibrations passed through the several selective devices, a common output for the vibrations from all of said selective devices, and a plurality of means for actuating said loudness controls in accordance with the characteristic syllabic changes of uttered speech for building up in said common output a speech-simulating composite of the vibrations so separately selected and controlled.

66. In a system for producing artificial speech, a source producing waves of frequencies extending throughout the essential speech range, a common output therefor, a plurality of means acting selectively throughout said frequency range for permitting waves of various frequency values to pass from said source into said common output intermittently and in any combination, means producing impulses corresponding to the

syllable rate of uttered speech, and means governed by said impulses for controlling said plurality of means to impress waves upon said common output in the sequences and combinations of uttered speech.

67. In a system for artificially producing speech, a source producing sound vibrations simulating the sustained vocal cord sounds, a source of vibrations simulating the unvoiced consonant sounds, means producing electrical impulses corresponding in frequency to the syllable rate of uttered speech, and means controlled by said impulses for releasing sound energy from said respective sources in proper sequence to simulate speech.

68. In a system for simulating speech, a source of harmonically related sound frequency vibrations, a source of continuous-spectrum vibrations, a common output channel for the vibrations from said sources, means operating at the syllable rate of uttered speech for controlling transmission from said sources into said channel, sound output means receiving its energy from said channel and a plurality of individual control devices associated with said channel each selectively and variably controlling transmission of vibrations from said channel to said sound output means in accordance with changes characteristic of speech.

69. In combination with a source of speech waves, means deriving therefrom in separate circuits code impulses of current of syllable rate or lower frequencies, means transmitting energy in accordance with said code impulses to a point of utilization, a source of sound frequency waves at said point, a plurality of selective devices for separating said sound frequency waves into small ranges for control purposes, means for individually controlling the amplitude of the waves comprised in said ranges in accordance with the received energy representing said code impulses, a

common energy output circuit, and means combining in said output circuit the waves so controlled from all of said selective devices.

70. The method of creating waves representing intelligible speech which comprises producing waves that lack information essential to intelligible speech, and artificially modulating them in accordance with waves that lack information essential to intelligible speech but contain the first-mentioned lacking information.

71. The method of creating waves representing intelligible speech which comprises producing electrical waves that lack information essential to intelligible speech, and modulating them with waves that lack information essential to intelligible speech but contain the first-mentioned lacking information.

72. The method of creating waves representing intelligible speech which comprises producing waves that lack information essential to intelligible speech, and modulating them by waves of syllabic frequency that lack information essential to intelligible speech but contain the first-mentioned lacking information.

73. The method of creating waves representing intelligible speech which comprises producing audio-frequency waves that lack information essential to intelligible speech, and artificially modifying them in accordance with waves that lack information essential to intelligible speech but contain the first-mentioned lacking information.

74. The method of creating waves representing intelligible speech which comprises producing audio-frequency electrical waves that contain a portion only of information essential to intelligible speech and modulating them by waves of syllabic frequency that lack information essential to intelligible speech but contain the information not contained in said produced waves that is essential to intelligible speech.

HOMER W. DUDLEY.