

ACOUSTIC APPARATUS FOR ENCODING SOUND

Filed Jan. 23, 1961

4 Sheets-Sheet 1

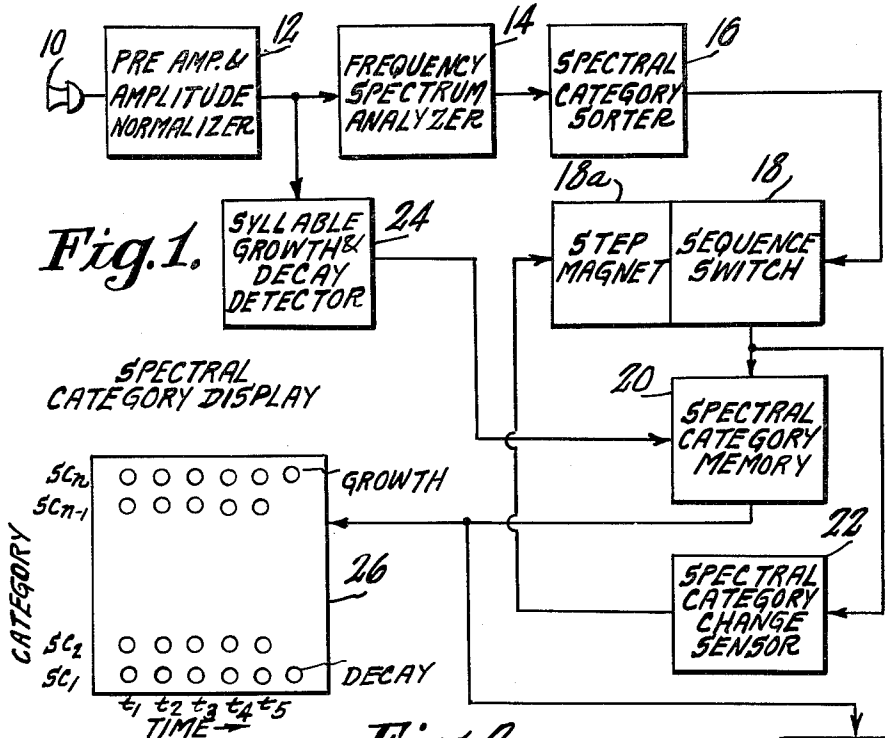


Fig. 1.

SPECTRAL CATEGORY DISPLAY

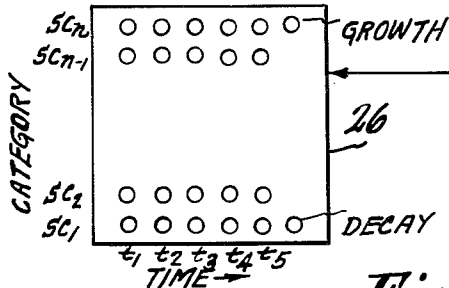


Fig. 2a.

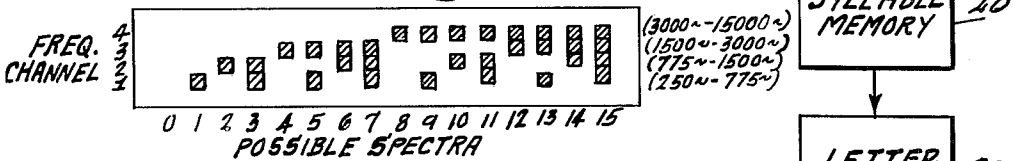


Fig. 2b.

SPECTRAL CATEGORIES					
SINGLE PEAK			DOUBLE PEAK		
LOCATION OF PEAK			LOCATION OF DIP		
L.F.	M.F.	H.F.	L.F.	M.F.	H.F.
CAT. NO.	CAT. NO.	CAT. NO.	CAT. NO.	CAT. NO.	CAT. NO.
1	2	3	4	5	6
1	2	8	5	9	10
3	4	12	13		11
7	6	14			

INVENTORS
 HARRY F. OLSON
 & HERBERT BELAR
 BY
 Morris A. Rabkin
 ATTORNEY

Aug. 31, 1965

H. F. OLSON ET AL

3,204,030

ACOUSTIC APPARATUS FOR ENCODING SOUND

Filed Jan. 23, 1961

4 Sheets-Sheet 2

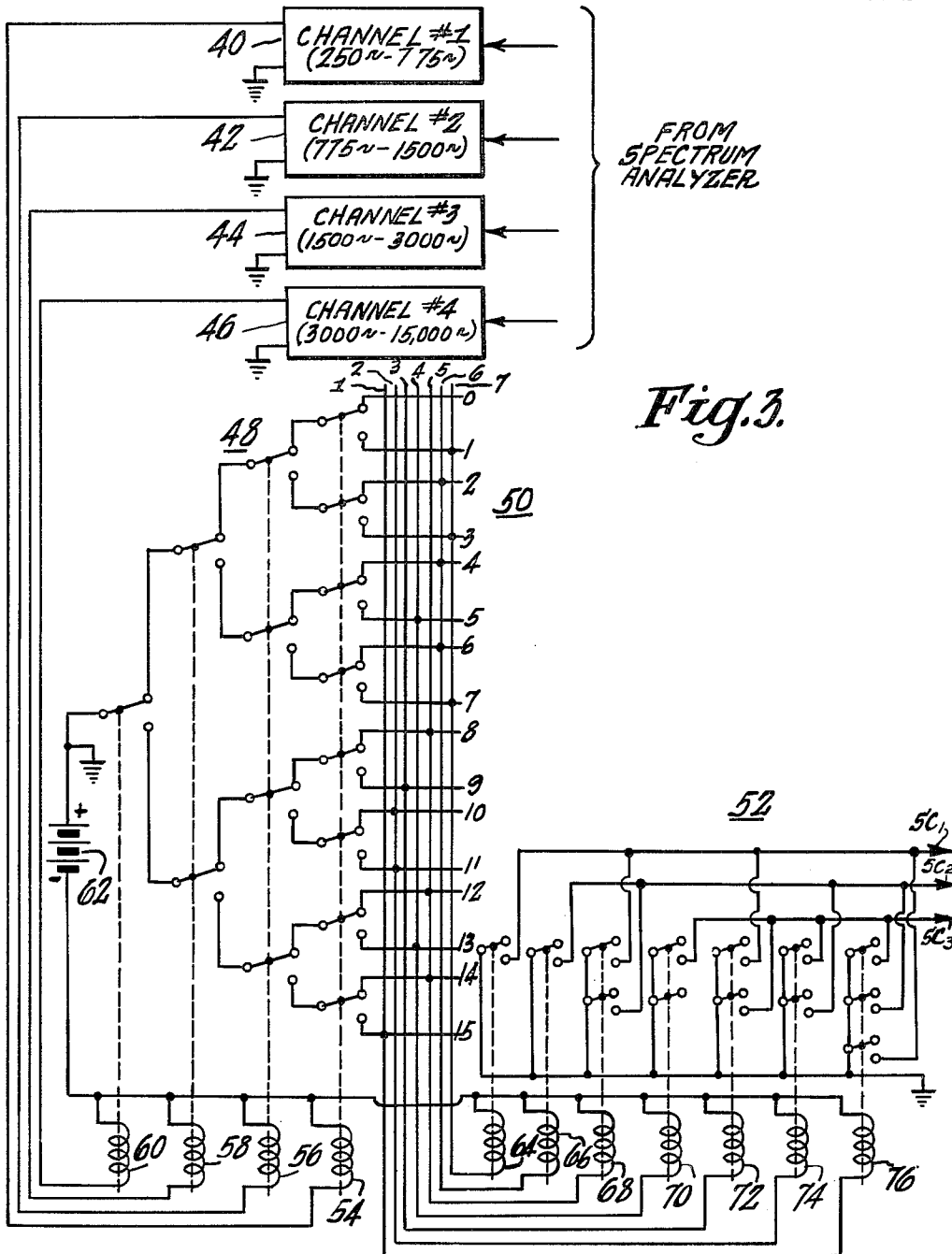
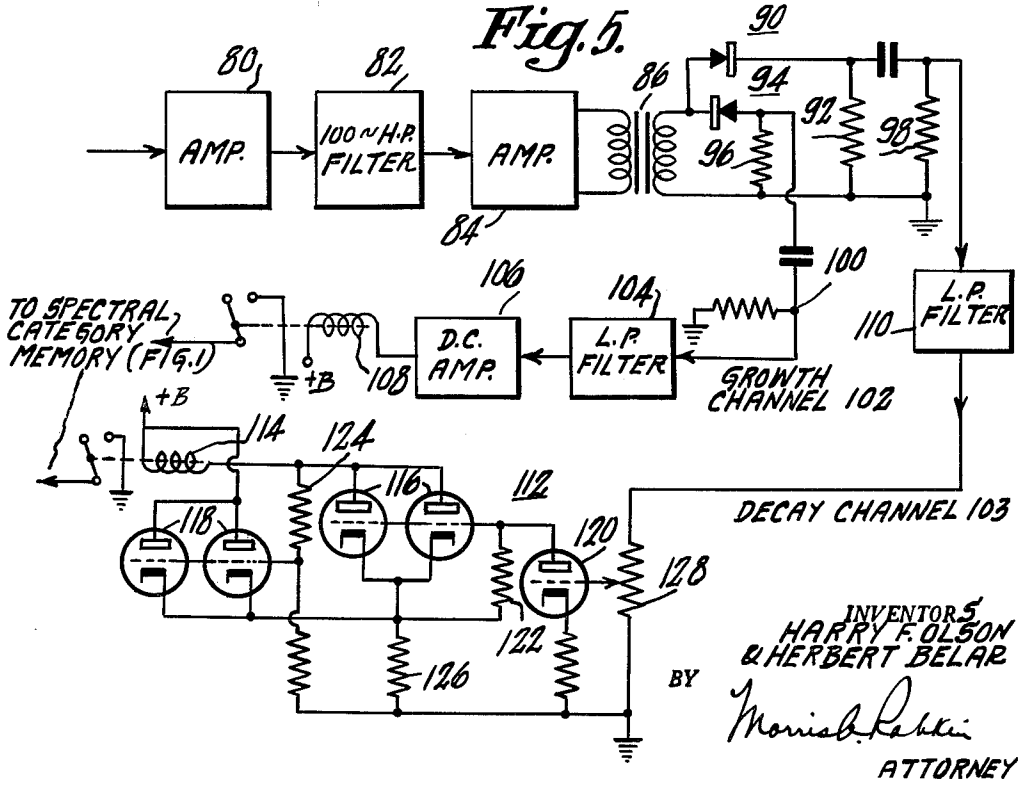
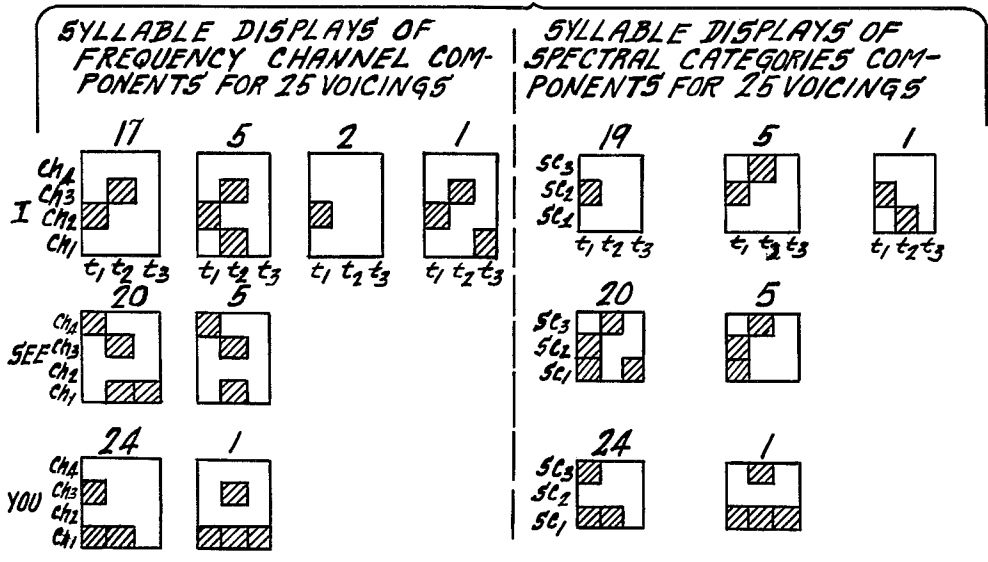


Fig. 3.

INVENTORS
HARRY F. OLSON
& HERBERT BELAR
BY
Morris Rabin
ATTORNEY

Fig. 4.



INVENTORS
 HARRY F. OLSON
 & HERBERT BELAR
 BY
 Morris B. Rabin
 ATTORNEY

Aug. 31, 1965

H. F. OLSON ET AL

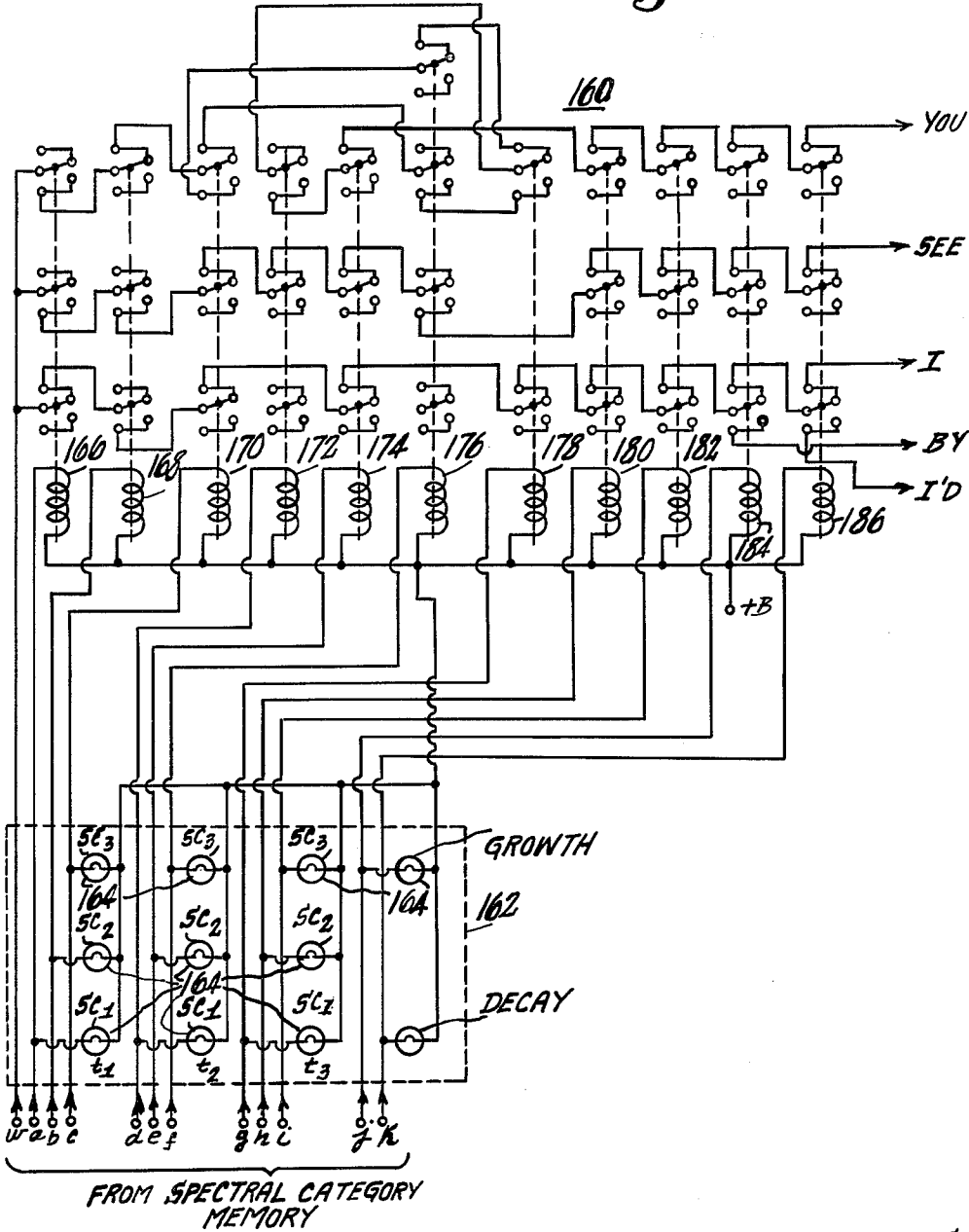
3,204,030

ACOUSTIC APPARATUS FOR ENCODING SOUND

Filed Jan. 23, 1961

4 Sheets-Sheet 4

Fig. 6.



INVENTORS
 HARRY F. OLSON
 & HERBERT BELAR
 BY *Morris Rabkin*
 ATTORNEY

1

2

3,204,030
**ACOUSTIC APPARATUS FOR ENCODING
SOUND**

Harry F. Olson, Princeton, and Herbert Belar, Palmyra,
N.J., assignors to Radio Corporation of America, a
corporation of Delaware

Filed Jan. 23, 1961, Ser. No. 84,229
20 Claims. (Cl. 179—1)

The present invention relates to acoustic apparatus, and more particularly to apparatus for recognizing speech and other sounds.

The invention is especially suitable for use in voice-operated apparatus for deriving signals for controlling the operation of a machine which automatically types or prints the speech spoken into the apparatus. Such voice-operated apparatus will be referred to hereinafter as a "phonetic typewriter." The invention is also suitable for use in apparatus for encoding and/or decoding sounds, such as speech, so that these sounds can be transmitted in the form of a code, such as a digital code. Since speech can be transmitted over a more limited bandwidth in the form of a digital code than would be the case when conventional modulation techniques are employed, the present invention is also generally useful in communications apparatus.

The present invention is an improvement upon the apparatus described and claimed in the following patents, applications for which were filed in the names of the present inventors: Patent No. 2,971,057, issued Feb. 7, 1961, Serial No. 490,592, filed Feb. 25, 1955, for Apparatus for Speech Analysis and Printer Control Mechanisms; and Patent No. 2,971,058, issued Feb. 7, 1961, Serial No. 662,370, filed May 29, 1957, for Method of and Apparatus for Speech Analysis and Printer Control Mechanisms. The subject matter of the foregoing patents may be found in: (1) a paper entitled "Phonetic Typewriter," which appears in the Journal of the Acoustical Society of America, Vol. 28, No. 6, November, 1956; and (2) a paper entitled "Time Compensation For Speed of Talking in Speech Recognition Machines," which appears in IRE Transactions on Audio, vol. AU-8, No. 3, May-June, 1960.

It is important for accuracy of analysis of sound to determine the characteristics of the formants which represent the sound. These formants describe the sound amplitude (energy) variations with frequency and they also describe the manner in which the frequency and amplitude of the sound vary with time. The formants are usually determined by dividing the frequency spectrum of sound into a multiplicity of frequency channels each carrying a different portion or band of the frequency spectrum. It is desirable to employ a large number of frequency channels in order to obtain the formant of the sound with greatest accuracy. The amount of information that is handled in completing the analysis of the sound is proportional to the number of frequency channels which are employed. It is desirable to facilitate the greater accuracy of analysis possible with a large number of frequency channels without appreciably increasing the complexity of the apparatus which handles the information obtained from the analysis and identifies the sounds; for example, in the case of speech sounds, as a word or syllable.

Other characteristics of formants of sound have a bearing upon sound analysis. It has been discovered that the growth and decay or rate of amplitude change with time of the sounds of speech syllables is a distinguishing characteristic of certain syllables. It is therefore desirable to utilize information as to the growth and decay of speech sounds in recognizing syllables.

The primary object of the present invention is therefore to provide improved apparatus for and method of analyzing sounds for automatic recognition of speech and other sounds.

It is a further object of the present invention to provide an improved phonetic typewriter.

It is a still further object of the present invention to provide improved apparatus for determining the formants of sounds with a high degree of accuracy.

It is a still further object of the present invention to provide improved automatic speech recognition apparatus which is less complex and lower in cost than such apparatus as has been heretofore suggested and yet maintains equal or better capacity for recognition of different speech syllables.

In a sound analyzing and speech recognition apparatus according to the present invention, the formants of the sounds to be recognized are obtained from a multiplicity of components each of which represents a different characteristic of the formants. For example, these components may be derived from different frequency channels in the frequency spectrum produced by the sound to be analyzed (hereinafter referred to as frequency channel or frequency spectrum components). Another component may be the growth of the sound as it is produced. Still another component may be the decay of the sound. The sound may be represented by a binary number including a plurality of binary bits each representing a different one of the components of the sound. The presence of the component, when it has an amplitude greater than a certain amplitude, will be indicated by a binary "one" bit, and the absence of the component will be indicated by a binary "zero" bit. It will be appreciated that the larger the number of components which enter into the analysis the more accurate the analysis. However, the number of possible different sounds represented by the components increases exponentially as powers of two in the present case, with the number of components which are considered in the analysis. The present invention includes means for sorting the combinations represented by a large number of components into a smaller number of categories or sets each containing selected combinations of components. These categories are encoded into a binary number having a smaller total of digits than the binary number representing all of the components upon which the analysis has been based. The categories may be mutually exclusive so that all of the information represented by all of the components is preserved. Alternatively, the categories may be selected on the basis of a limited vocabulary of sounds, such, for example, as a certain number of speech syllables. In the latter case, those sounds which correspond to the same speech syllables, although they may have different combinations of components because they originate with different speakers, are allotted to the same category.

Automatic recognition of the speech may be carried forward in the same manner as described in the referenced patents and articles by the present inventors. However, the categories of the components, rather than the components themselves, provide the basis for the recognition. The apparatus described in the referenced publications can operate upon the binary numbers which represent the categories with equal facility as it can operate upon the binary numbers which represent the components. Accordingly, the effective capacity of an automatic sound recognition apparatus, which ordinarily operates upon the components of sound, may be increased, since categories are utilized, each containing a large number of component combinations, rather than the components themselves.

The invention itself, both as to its organization and

3

method of operation, as well as additional objects and advantages thereof, will become more readily apparent from the following description when read in connection with the accompanying drawings in which:

FIGURE 1 is a schematic, block diagram of a system for analyzing the sound of speech syllables in accordance with an illustrative embodiment of the present invention;

FIGURE 2a is a chart showing possible combinations of frequency components in spectra which may be obtained in a system similar to the system shown in FIG. 1, but which is simplified for purposes of illustration;

FIGURE 2b is a schedule of categories of the possible spectra which are illustrated in FIG. 2a;

FIGURE 3 is a schematic diagram, partially in block form, of a spectral category sorter of the type illustrated in block form in FIG. 1, but which is simplified for the purpose of illustration;

FIGURE 4 is a pair of charts showing syllable displays of frequency spectrum components and their corresponding spectral category components;

FIGURE 5 is a schematic diagram, partially in block form, of a syllable growth and decay detector of the type shown in block form in FIG. 1; and

FIGURE 6 is a schematic diagram, partially in block form, illustrating the operation of a spectral category display and of a syllable memory of the type which are shown in block form in FIG. 1, but which is simplified for purposes of illustration.

Referring more particularly to FIG. 1 of the drawings, there is shown a microphone 10 for translating speech into corresponding electrical signals. These electrical signals are applied to a pre-amplifier and amplitude normalizer 12. The amplitude normalizer may include circuits which are known in the art for maintaining a relatively constant output signal level over a wide range of input signal levels. Since the analysis of the sounds of speech is based upon the amplitude of the speech signals derived by the microphone 10, it is desirable to compensate for the variations in amplitude level which ordinarily occur in the course of normal speech.

Following amplitude normalization, the speech signals are applied to a frequency spectrum analyzer 14. This analyzer includes a multiplicity of frequency channels each including a frequency selective network which passes only a selected portion or band of the speech frequency spectrum. Each frequency channel also includes an amplifier followed by a rectifier which translates the output of the frequency channel into a direct current signal having an amplitude dependent upon the amplitude of the speech signal in the band of the spectrum covered by the channel. A direct current amplifier is provided in each channel for amplifying the direct current signal. The design of the frequency spectrum analyzer is described in greater detail in the above referenced article which appears in the Journal of the Acoustical Society of America and also in the above referenced patents.

Each frequency channel provides a different component of the speech signals and therefore of the speech itself. The number of components depends upon the number of channels in the frequency spectrum analyzer. For example, if sixteen different frequency channels are incorporated in the analyzer 14 and each component may be considered to be present or absent if its amplitude exceeds a predetermined threshold level, there may be 2^{16} , or 65,536 different possible combinations of components including the combination in which no components are present. Assuming that the frequency spectrum of these components is sampled during five discrete time steps for each syllable, as is the case in the systems described in the above referenced patents and publications, the possible combinations of components which might be present for each syllable is 2^{80} , or several trillion possible combinations of components. Analysis taking into account all of the possible combinations becomes exceedingly difficult, if not impossible. It is a feature of the

4

invention to provide means for reducing the complexity of the analysis while retaining the significant information obtained from the frequency spectrum analysis.

A spectral category sorter 16 in accordance with a feature of the invention provides means for reducing the complexity of the analysis. This sorter is a switching system which automatically allocates selected combinations of frequency components provided by the frequency spectrum analyzer 14 into categories or sets of spectra in accordance with a program which is built into the sorter 16. The sorter will be described in detail hereinafter in connection with FIG. 3 of the drawings. Included in the switching system of the sorter may be amplitude sensitive devices, such as threshold relays, for converting the frequency spectrum components into voltages representing binary numbers "one" or "zero" depending upon whether the amplitude level of the component exceeds or does not exceed a predetermined threshold level, respectively. This threshold level is determined (1) by the operation of the amplitude normalizer 12 and (2) by the voltage drops in the spectrum analyzer 14, which also takes into account the sensitivity of the relays in the spectral category sorter 16.

The outputs of the spectral category sorter are spectral category components later identified as SC_1 to SC_n . These are distributed through a sequence switch 18 to a spectral category memory 20 and to a spectral category change sensor 22. The sequence switch may be a telephone type stepper switch having a plurality of wipers one for each category and a plurality of successive levels of contacts each including terminals corresponding in number to the number of wipers. Each level of contacts provides a succeeding time step. The sequence switch 13 is described in the above referenced publications and patents. In the case of the system described in these publications, a sufficient number of levels of contacts is provided to accommodate five time steps or intervals.

The spectral category change sensor 22 is a switching circuit that controls the operation of a step magnet 18a which is incorporated in the sequence switch 18. The spectral category change sensor 22 responds to changes in the combinations of spectral categories which are obtained from the spectral category sorter much in the same manner as the frequency component sensor described in the above referenced IRE Transactions publication and in the above referenced Patent No. 2,971,058. The step magnet 18a is not actuated and the sequence switch wipers remain on the first time step terminals until one or more of the spectral category components change. When a change occurs, the step magnet 18a is energized and the sequence switch will move to the next time step terminals. When the sounding of a syllable is completed, the sequence switch returns to its start position. Since the spectral category change sensor operates the sequence switch 18 only upon a change in the spectral category components, the time intervals at which the spectral category components are analyzed will vary with the speed of talking. The system, as shown in FIG. 1, is therefore compensated for the speed of talking, as follows from the fact that the frequency spectrum components of speech will vary only when significant changes in the speech spectrum occur. The spectral category components correspond to the significant frequency spectrum components and vary only when significant changes in the speech spectrum occur. The speed of talking is compensated since the spectral category change sensor will follow the changes in the speech spectrum.

The spectral category memory 20 provides storage for binary numbers representing each of the spectral category components at each time step established by the sequence switch 18 during which analysis of a speech syllable takes place. The spectral category memory 20 may be the same as the spectral memory described in the above referenced publications and patents in that it

5

will contain a relay for each spectral category in each time step during which the spectral categories are sampled by the sequence switch. These relays are energizable by the signal outputs from the spectral category sorter 16 which are transferred into the spectral category memory 20 by way of the sequence switch 18. In the spectral memory described in the referenced publications and patents, the relays of the spectral memory each provide storage for a different frequency spectrum component at each time step. The spectral category memory also has capacity for storage of components representing the growth and decay of a speech syllable. Storage may be provided by a pair of relays which are additional to the relays which are provided for the spectral category components.

The growth and decay components are obtained for storage in the spectral category memory by a syllable growth and decay detector 24. The signals from the preamplifier and amplitude normalizer 12 are applied to the syllable growth and decay detector 24. This detector 24 will be described in greater detail hereinafter in connection with FIG. 5 of the drawings. Briefly, it includes a pair of channels, one for growth and the other for decay, which respond separately to the rate of change of amplitude of a signal representing a speech syllable in a positive direction and in a negative direction. A signal corresponding to the growth of the speech syllable signal is translated into a voltage representing a binary "one" bit when it exceeds a predetermined signal level and a binary "zero" when it does not exceed this predetermined signal level. The signal level is a function of the rate of growth. Similarly, the signal representing the decay of the speech syllable signal is translated into a binary "one" bit or into a binary "zero" bit depending upon the rate of the decay.

The information stored in the spectral category memory 20 is displayed on a spectral category display 26. This display includes a plurality of lights arranged in rows each representing a different spectral category component SC_1 to SC_n , the exact number of components depending upon the number of spectral categories into which the frequency spectrum components are sorted in the sorter 16. The columns represent different time steps during which the spectrum categories are sampled by virtue of the operation of the sequence switch 18. Lights may be provided to represent the growth and decay components which are stored in the spectral category memory 20. The illumination of a light may indicate the presence of a spectral category component or of growth and decay at greater than a predetermined rate.

The output of the spectral category memory 20 is transferred at the end of the time steps (five time steps being illustrated in FIG. 1) to a syllable memory 28. This syllable memory 28 may be similar to the syllable memory which is described in the above referenced publications and patents except that the syllables stored therein are related to different combinations of spectral categories, rather than to different combinations of frequency spectrum components. The operation of the syllable memory 28 will be more apparent from FIG. 6 of the drawings which will be described hereinafter. Briefly described, the syllable memory includes a switching system comprising a plurality of relays having their contacts connected in accordance with binary codes to store a different syllable for each code. The relays in the syllable memory 28 are selectively energized by different combinations of spectral category components and growth and decay components which are stored in the spectral category memory 20. When, at the end of the sounding of a syllable, the binary number stored in the spectral category memory 20 is transferred to the syllable memory 28, and the transferred binary number corresponds to a previously mentioned syllable memory code, the syllable corresponding to the category com-

6

ponents stored in the spectral category memory 20 is transferred to and stored in the syllable memory 28.

The information in the syllable memory is transferred to a letter decoder 30. This decoder is described in the referenced patents and publications. It includes a rotary switch and a matrix which establishes a plurality of different connections from the contacts of the relays in the syllable memory 28 through the rotary switch. Each of these connections may correspond to a different letter of the syllables stored in the syllable memory 28. Thus, an automatic typewriter 32 may be controlled by the decoder 30 to print out the letters of the syllable stored in the syllable memory 28 which corresponds to the binary number stored in the spectral category memory 20. The typewriter will print out sequences of syllables which can be understood by the reader. The output of the syllable memory may alternatively be connected to a digital information transmission system, either wire or wireless, for the purpose of communicating the speech information in digital form. The different binary numbers will be obtainable at the output of the syllable memory depending upon which syllable is stored in the syllable memory.

Referring to FIG. 2a, all possible spectra which are produced with a frequency analyzer having four frequency channels are shown. Each frequency channel component is considered "present" or "absent" depending upon whether its amplitude exceeds or does not exceed a predetermined signal level threshold. The frequency channel components which are present are indicated in FIG. 2a as blocks which are filled with hatching. When a component is "absent," it is represented by a blank region in FIG. 2a. The possible spectra may therefore be presented in accordance with a binary code where the "present" condition of a component is represented by a binary "one" bit and the "absent" condition of a component is represented by a binary "zero" bit. Since there are four frequency channels, there are sixteen possible combinations of frequency components, or sixteen possible spectra.

Each of the possible spectra may represent a different formant of a speech sound. The formants are usually characterized by the number and location (frequency wise) of amplitude peaks. When the formants are represented by frequency channel components which are displayed in binary form, a peak may be represented by a single one or a cluster of adjacent "present" or binary "one" frequency channel components. The center of the peak may be considered to be in the frequency channel of the centrally located one of a cluster of frequency channel components. It is desirable that a large number of frequency channels be used so that the peaks in the formants can be located with precision. In the speech analysis apparatus described in the above referenced publications and patent applications, eight frequency channels are used. It is desirable to use even more than eight frequency channels. Experience indicates that as many as one hundred frequency channels would be desirable in providing a precise analysis of the formants resulting from the sounds of speech. If a binary code is utilized, the number of frequency channels is desirably some power of two. Thus, a suitable frequency spectrum analyzer may include one hundred and twenty-eight frequency channels (2^7 channels). It should be understood, however, that the number of frequency channels may be smaller or greater than the aforementioned number. For purposes of the present description, only four frequency channels are shown in order to simplify the illustration. It will be observed, however, that the principles of analysis and word recognition are the same regardless of the number of frequency channels involved.

Four frequency channels result in sixteen possible spectra. The number of possible spectra increases by a factor of two for each additional frequency channel. It, therefore, becomes extremely difficult to handle all of

the information which results from the frequency spectrum analysis when a large number of frequency channels is employed in the frequency spectrum analyzer. It is a feature of the present invention to resolve the possible spectra into spectral categories, or sets of frequency channels, as they appear in their possible spectra. These spectral categories may be established empirically, or in accordance with some other predetermined relationship between the possible spectra and the spectral categories.

FIG. 2b is a schedule of spectral categories which are established in accordance with a predetermined relationship between possible spectra. This relationship is derived from the formants which are represented by the various possible spectra. These formants may have a single peak or a double peak. A peak may be located in any one of the four frequency channels. A dip appears between each of the double peak formants. This dip may be also located in the frequency channels. Six categories (cat. No. 1 to cat. No. 6) are established. Categories Nos. 1 to 3, inclusive, include those spectra having a single peak. Categories Nos. 4 to 6, inclusive, include those spectra having a double peak. The location of a single peak in the low frequency channel (250 cycles to 775 cycles), mid frequency channels (775 to cycles to 3,000 cycles), or high frequency channels (3,000 cycles to 15,000 cycles) determines whether the spectra will be in the first, second or third category, respectively. Similarly, the location of the dip in a double peak spectrum in the low frequency channel, mid frequency channels or high frequency channels will determine whether the spectrum is categorized in the fourth, fifth or sixth category, respectively. Two additional categories not shown on the schedule FIG. 2b are included. Category No. 0 is for the case where all of the frequency channels are absent and are represented by binary "zero" bits. Category No. 7 is provided for the all response spectrum, which is number fifteen in FIG. 2a, where all of the frequency channels are present and are represented by binary "one" bits.

By utilizing spectral categories which are established in accordance with a predetermined relationship among the various possible spectra, which in the aforementioned case is the location of peaks and dips in the formants, the number of possible representations of the sound information has been reduced. The reduction is from sixteen representations to eight representations in the illustrated case. The total amount of information in the possible spectra has been condensed into the spectral categories. The various spectral categories are mutually exclusive and the information and capacity for speech recognition has been preserved with a smaller number of information elements. Thus, the storage capacity of storage devices in a phonetic typewriter, such as the spectral category memory 20 in FIG. 1 need only be a relatively small number of memory units, when the apparatus is adapted to handle spectral categories.

If the spectral categories were allocated and established in accordance with some routine or program which was arrived at empirically, a still smaller number of spectral categories might be possible. It will be more desirable to use empirically arrived at categories when a large number of frequency channels, say, over one hundred, is employed. One empirical basis upon which the spectral categories may be established may be the frequency of occurrence of the same spectra in response to voicing of a certain sound. For example, the sound "you" may produce, during the first of five intervals of time during which it is sounded, the same ten of one hundred and twenty-eight possible spectra for ninety out of one hundred voicings. A category may be established for these ten recurring spectra. In a similar manner, a group of categories may be established for the sounds of a selected number of speech syllables which are to be recognized by the apparatus.

By proceeding systematically through the vocabulary

of syllables which are selected for automatic recognition, and by repeated voicings of each of the syllables, a group of spectral categories may be empirically established. These spectral categories may not include every possible spectrum which might be produced by a multi-channel frequency spectrum analyzer. All those spectra which are significant, in that they correspond to certain of the selected syllables, will, however, be represented in various ones of these spectral categories. Once the spectral categories are determined on an empirical basis, these spectral categories may be built into a spectral category sorter, such as the sorter 16 shown in FIG. 1.

A spectral category sorter into which the spectral categories shown in the schedule of FIG. 2b are built is shown in FIG. 3. The frequency channels Nos. 1, 2, 3 and 4 in the frequency spectrum analyzer each include a rectifier and a direct current amplifier in the output stages thereof. When the signals in the various frequency channels are above a predetermined threshold level, they will bias the direct current amplifier associated with the respective channels to conduct sufficient current to operate a threshold relay and cause it to pull in. The sorter, itself, is a switching system which includes a relay tree 48, a matrix of conductors 50, and another relay tree 52. The relay tree 48 provides sixteen circuit paths, each corresponding to a different one of the sixteen possible spectra shown in FIG. 2a. The operating windings of the relays 54, 56, 58 and 60 of the relay tree 48 are connected to the outputs of the frequency channels Nos. 1, 2, 3 and 4, respectively. When a signal of a level greater than the predetermined level is present in any frequency channel, current from a source of operating voltage, illustrated herein as a battery 62, passes through the relay operating winding 54, 56, 58 or 60 associated with that channel and causes the relay to pull in.

The matrix of conductors 50 may be a plug board having a plurality of intersecting conductors and plugs for making connections at selected intersections of the conductors. The plugs are shown as heavy dots in FIG. 3. Sixteen conductors, which correspond to the sixteen possible spectra, are disposed horizontally in the matrix 50 and seven conductors, which correspond to categories Nos. 1 to 7, are vertically disposed. Each of the horizontally disposed conductors is intersected at seven points. Each of the sixteen horizontal conductors in the matrix may be connected to ground through the switch contacts of the relay tree 48. Each of the seven vertical conductors is connected to a different operating winding of the relays 64, 66, 68, 70, 72, 74 and 76 of the relay tree 52. The relays 64, 66, 68, 70, 72 and 74 correspond to categories Nos. 1 to 6, respectively, on the schedule of FIG. 2b. The relay operating winding 76 corresponds to category No. 7, or the condition that frequency channel components are present in all of the channels 40, 42, 44 and 46. The category No. 0 is represented by the absence of a response in any of the frequency channels, so that none of the relays 64, 66, 68, 70, 72, 74 and 76, operates upon occurrence of category No. 0.

The matrix 50 is wired in accordance with schedule of FIG. 2b. The operating windings of the relays 64, 66, 68, 70, 72, 74 and 76 are connected to the high potential side of the battery 62, and can be selectively connected to ground through the matrix 50 and the switch contacts of the relay tree 48. The relay tree 52 translates categories Nos. 1 to 7 into three spectral category components SC₁, SC₂ and SC₃. Different combinations of these components correspond to different ones of the categories Nos. 1 to 7.

The operation of the spectral category sorter will be apparent from the following example: When an output in frequency channels 2 and 3 occurs, spectrum No. 6 of the possible spectra illustrated in FIG. 2a occurs. Relays numbers 56 and 58 will pull in, thereby connecting ground to the horizontal conductor 6 in the matrix 50.

The operating winding of relay 66 in the relay tree 52 is then connected to ground through a circuit including the horizontal conductor 6 of the matrix 50. Relay 66 will pull in, thereby connecting ground to the spectral category component SC_2 output. The spectral category component outputs SC_1 , SC_2 and SC_3 are connected to the sequence switch 18 (FIG. 1). Assuming that the sequence switch is a stepper switch, as mentioned above, the outputs of the respective spectral category components SC_1 , SC_2 and SC_3 will be connected separately to different ones of the wiper arms of the switch. A binary number of three bits corresponding to the outputs of the components SC_1 , SC_2 and SC_3 will be stored at a plurality of time intervals in the spectral category memory 20. When an output terminal for components SC_1 , SC_2 or SC_3 is connected to ground, the corresponding spectral category component will be stored in the spectral category memory 20 as a binary "one" bit, and when ungrounded as a binary "zero" bit. Thus, three binary bits representing the eight spectral categories will be sufficient to indicate the absence or presence of the spectral categories in the spectral category memory 20.

Apparatus in accordance with the invention, as described above, therefore includes means which separate the sound into a plurality of components each of which are translated into binary digits. These digits constitute a binary number representing the components. This binary number is compared with defined binary numbers which correspond to different categories of components. The components are sorted into these categories, when the binary number which represents the components corresponds to any one of the defined binary numbers for the categories. The categories themselves are represented by discrete binary digits which in turn form a binary number. The category binary number has fewer digits than the component binary number. Nevertheless, the category binary number represents the sound.

Referring to FIG. 4, displays of spectral category components for 25 voicings, each of the syllables "I," "see" and "you" are shown. It will be noted that the syllable displays of four frequency channel components Ch_1 , Ch_2 , Ch_3 and Ch_4 are illustrated and three time steps t_1 , t_2 and t_3 , while the syllable displays are for three category components SC_1 , SC_2 and SC_3 which appear at the three time steps t_1 , t_2 and t_3 . The number of displays out of 25 voicings which are the same for each of the syllables "I," "see" and "you" in the case of the frequency channel components are also the same in the case of the spectral category components. It will be noted that, in the case of the syllable "I," nineteen out of 25 displays of category components are the same, whereas only seventeen out of twenty-five displays are the same in the case of the frequency channel components. The wiring in the spectral category memory 20 and in the syllable memory which stores different enunciation of the syllable "I" may therefore be simplified, since wiring to accommodate one voicing of the syllable "I" may be omitted. It will be further noted that the storage capacity of the spectral category memory 20 may be much smaller than the storage capacity of a spectral memory which serves the similar purpose in the referenced patents and publications and yet accommodates the same number of syllables. There are twelve bits and eight possible combinations of these bits to recognize the syllables "I," "see" and "you," or ninety-six total bits storage capacity required in a spectral memory of the known type. In the spectral category memory 20, on the other hand, there are nine bits and seven possible combinations of these nine bits to recognize the syllables "I," "see" and "you," resulting in a total of sixty-three bits as compared to ninety-six bits for the syllable memory which stores the frequency channel components. A similar or greater reduction in storage capacity results from the use of the present invention in a syllable recognition system having capacity to utilize a large number of frequency channels and spectrum cate-

gories, and which has the capacity for recognizing a larger vocabulary of syllables.

Another component which is of significance in the recognition of syllables is the growth and decay of the syllable as it is sounded. A detector which is suitable for detecting the growth and decay of a syllable and provides a growth component, if the growth of the syllable is greater than a certain rate, and a decay component, if the syllable decays at greater than a certain rate, is shown in FIG. 5 of the drawings.

In addition to the growth and decay characteristics of a syllable, there are other characteristics of speech sounds which have been found useful in providing components which enter into the analysis of sounds and particularly the sounds of speech. The use of these components makes the recognition of certain speech sounds more precise. However, the equipment for analyzing the sound and discriminating among different syllables may become more complex when these additional components are considered. For example, one component which is useful in analyzing certain speech sounds is the fundamental frequency of the speech. This fundamental frequency is usually below 200 cycles per second and is usually suppressed by inserting a high pass filter in the preamplifier, such as the preamplifier 12 (FIG. 1). For example, those syllables containing the sounds represented by the symbol "e" and "i" of the international phonetic alphabet may be difficult to distinguish unless the fundamental frequency of the sounds of these syllables enters into the analysis as a separate component thereof. The syllable represented by "e" is used in the word red, whereas the syllable represented by the symbol "i" occurs in the word read. Thus, if the fundamental frequency is greater than a certain frequency, a spectrum may be sorted into a category representing the syllable "e," whereas if the fundamental frequency is below a certain frequency and the same spectrum results, this spectrum would be sorted into a category representing the syllable "i." Thus, in addition to the frequency channel components which enter into the spectral category, fundamental frequency components may also be provided which will be taken together with the frequency channel components in determining whether certain of these frequency channel components should be allocated to different ones of the pre-established spectral categories.

Returning, now, to the growth and decay detector shown in FIG. 5 of the drawings, the signals from the preamplifier and amplitude normalizer 12 (FIG. 1) are amplified in an audio amplifier 80 (FIG. 5) and filtered after amplification in a high-pass filter 82 having a cut-off frequency of 100 cycles per second. This high-pass filter 82 removes the rumble or low frequency noise, which might be picked up by the microphone 10. The signal output is amplified in another audio amplifier 84 which may have a push-pull output stage including an output transformer 86. The output from the transformer 86 is rectified by a rectifier circuit 90 which delivers a positive (with respect to ground) output voltage across a resistor 92 and a second rectifier circuit 94 which delivers a negative (with respect to ground) output voltage across an output resistor 96. The output voltage of the rectifier circuit 90 is fed to a differentiating circuit 98, and the output of the other rectifier circuit 94 is fed to a similar differentiating circuit 100. The differentiating circuit 98 provides a negative signal level in response to decay of the signal voltage across the output of the resistor 92. The rectifier circuit 100 provides a negative output signal voltage in response to the growth of the output voltage across the resistor 96 in the rectifier circuit 94. The rectifying circuit 94 and the differentiating circuit 100 are included in a growth channel 102 together with a low pass filter 104 and a direct current amplifier 106, the latter of which operates a relay 108. The rectifier circuit 90 and the differentiating circuit 98 are included in a decay channel 103 together with a low pass filter 110 which is similar

11

to the filter 104, and a direct current amplifier 112 and a relay 114 which is operated by the amplifier 112.

The direct current amplifiers 106 and 112 may be substantially identical, the circuit of only the amplifier 112 being shown by way of example. The amplifier 112 includes a pair of relay control tubes 116 and a pair of load control tubes 118. Each of these pairs of tubes can be replaced by a single tube having sufficient plate current carrying capacity. When a negative voltage is applied to the grid of an input tube 120 of the amplifier 112, the voltage across its load resistor 122 becomes positive and a positive voltage appears at the grids of the tubes 116. The tubes 116 are normally cut off since the tubes 118 are biased into conduction by a positive voltage which appears across a voltage divider 124. With the tubes 118 in their conductive state, insufficient current flows through the winding of the relay 114 to cause the relay 114 to pull in, since the voltage drop across the voltage divider 124 leaves insufficient voltage across the winding of the relay 114 to develop current to pull in the relay contacts. However, when the tubes 116 conduct due to the application of sufficiently negative voltage to the grid of the input tube 120, the voltage drop across the resistor 122 in the plate current path in the input tube 120 decreases. A positive voltage then appears on the grids of the tubes 116. The tubes 116 conduct and the voltage drop across their cathode resistor 126 increases. The voltage at the cathodes of the tubes 118 rises with respect to the voltage at the grids of the tubes 118, thereby causing the tubes 118 to cut off. Current now flows through the operating winding of the relay 114 by way of the plate-cathode paths through the tubes 116 which is of sufficient magnitude to cause the relay 114 to pull in. The threshold level for relay pull-in may be adjusted by a potentiometer 128 which is connected to the grid of the input tube 120. The contacts of the relay 114 may be connected to a relay in the spectral category memory 20 (FIG. 1). When the relay 114 pulls in, a binary "one" bit, indicating that the decay rate of the syllable is greater than a certain rate, is stored in the spectral category memory.

The relay 108 operates in manner similar to the relay 114. Accordingly, when the growth rate of the syllable is greater than a certain rate, the relay 108 will pull in. A relay in the spectral category memory 20 operates, upon pull-in of the relay 108, and stores a binary "one" bit in the spectral category memory, indicating that the syllable growth rate is greater than the predetermined rate previously established by the design of the growth and decay detector.

The rectifier circuit 90 and the differentiating circuit 98 operate as follows to provide a negative voltage level in response to the decay of a syllable: The voltage across the output of the transformer 86 varies both in a positive and a negative direction with respect to ground. Only the positive going portion of the transformer output voltage passes through the rectifier circuit 90 and appears across the resistor 92. The slope of the envelope of this voltage is positive going at the beginning of a syllable and negative going at the end of the syllable. Thus, when the syllable decays, at the end thereof, the envelope decreases at a rate depending upon the rate of decay. The differentiating circuit 98 differentiates the envelope and, since the envelope has a negative slope, produces a negative voltage in response thereto. This negative voltage is filtered in the filter 110 and is applied across the potentiometer 128. The amplitude of this negative voltage depends upon the rate of the decay. Accordingly, when the decay exceeds its predetermined rate, the direct current amplifier 112 will operate to cause the relay 114 to pull in and register a binary "one" bit indicative of a decay component in the spectral category memory 20 (FIG. 1). The growth portion of the envelope rectified by the rectifier 90 and differentiated by the differentiating circuit 98 has no effect in the decay channel since it is a

12

positive going voltage which has no effect on the amplifier 112 and does not change its state of operation.

The rectifier circuit 94 in the growth channel passes only that portion of the voltage output of the transformer 86 which is negative going with respect to ground. The leading edge of the voltage envelope representing a syllable increases in the negative direction as the syllable grows in amplitude. The differentiating circuit 100 provides a negative voltage in response to this negative going growth signal which has a level depending upon the rate of growth. This negative voltage is filtered and causes the direct current amplifier 106 to operate the relay 108 when the growth rate of the syllable is greater than the predetermined rate. Accordingly, the relay 108 will operate to cause a binary "one" bit to be stored in the proper position in the spectral category memory 20.

Referring, now, to FIG. 6 of the drawings, there is shown a simplified syllable memory which recognizes syllables based upon the information stored in the spectral category memory 20. The illustrated syllable memory is capable of handling the syllables "I," "see," and "you" and the syllables "by" and "Td." It will be noted that the syllables "I," "see" and "you" are those which were taken by way of example in illustrating the mode of operation of the spectral category sorter 16, FIG. 1. The syllable memory shown in FIG. 6 also has capacity for storing the syllables "Td" and "by" in order to illustrate the capability of the decay and growth components of the speech which are stored in the spectral category memory 20.

FIG. 6 illustrates a simplified syllable memory 160 and a spectral category display 162 which displays the spectral category components SC_1 , SC_2 and SC_3 which are stored in the spectral category memory after three intervals of time t_1 , t_2 and t_3 . The growth component and the decay component are also displayed in the display 162. A lamp 164 is included in the display 162 for each of the spectral category components and for the growth and decay components. Each of the spectral category components and the growth and decay component is either present or absent in the spectral category memory. Accordingly, a different binary bit a to k , inclusive, may be considered to correspond to each of these spectral category components and the growth and decay components. Thus, binary bits a , b and c correspond to the spectral category components SC_1 , SC_2 and SC_3 which are stored in the spectral category memory at time t_1 . Binary bits d , e and f correspond to the spectral category components SC_1 , SC_2 and SC_3 which are stored in the spectral category memory at time t_2 , respectively. Binary bits g , h and i correspond to the spectral category components SC_1 , SC_2 and SC_3 which are stored in the spectral category memory at t_3 . The binary bit j corresponds to the growth component and the binary bit k corresponds to the decay component. The various displays for the syllables "I," "see" and "you" which would be obtained with the spectral category display 162 have previously been described with reference to FIG. 4 of the drawings. The presence or absence of the growth and decay components is used to distinguish the syllable "I" from the syllables "by" and "Td."

The syllable memory includes 11 relays 166, 168, 170, 172, 174, 176, 178, 180, 182, 184 and 186. Each of these relays has a plurality of sets of contacts. The set of relay contacts which is shown closest to the operating winding of the relay will be called the first relay contact set. The set of relay contacts next closest from the operating winding will be called the second relay contact set and so on in accordance with their positional relationship. The first relay contact sets of the relays 166 to 186 are wired to store either the syllable "I" or the syllable "by" or the syllable "Td." The second relay contact sets of all of the relays are wired to store the syllable "see." The third and fourth relay contact sets are wired to store the syllable "you." The wiring of the

sets of relay contacts is done in accordance with syllable recognition codes for the syllables "I," "Pd," "by," "see" and "you." These codes are determined by the spectral category components of the syllables which, in the case of the syllables "I," "see" and "you," are displayed in FIG. 4. The syllable recognition codes for all of the syllables which are stored in the memory 160 are shown in the following table:

SYLLABLE RECOGNITION CODES

t ₁			t ₂			t ₃			Gr	Dc	Syllable
SC ₁	SC ₂	SC ₃	SC ₁	SC ₂	SC ₃	SC ₁	SC ₂	SC ₃			
a	b	c	d	e	f	g	h	i	j	k	Syllable
0	1	0	0	0	0	0	0	0	0	0	I
0	1	0	0	0	1	0	0	0	0	0	I
0	1	0	1	0	0	0	0	0	0	0	I
0	1	0	0	0	0	0	0	0	0	1	Pd
0	1	0	0	0	1	0	0	0	0	1	Pd
0	1	0	1	0	0	0	0	0	0	1	Pd
0	1	0	0	0	0	0	0	0	1	0	By
0	1	0	0	0	1	0	0	0	1	0	By
0	1	0	1	0	0	0	0	0	1	0	By
1	1	0	0	0	1	0	0	0	0	0	See
1	1	0	0	0	1	1	0	0	0	0	See
1	0	1	1	0	0	0	0	0	0	0	You
1	0	0	1	0	1	1	0	0	0	0	You

It will be observed, from the table, that different binary numbers including the bits a to k identify the syllables "I," "Pd," "by," "see" and "you." It will further be noted that the binary numbers for the syllables "I" and "Pd" differ only in the presence of a binary "one" bit k. The binary numbers for the syllables "I" and "by" differ only by the presence of a binary "one" j bit. Since the j bit is present when the growth of the syllable is faster than a certain rate and a binary "one" k bit is present when the decay of a syllable is greater than a certain rate, the growth and decay components singularly determine the difference between the syllables "I," "by" and "Pd."

A line of lead w in FIG. 6 is connected to the tongue of the contact sets associated with the relay 166. Either ground or a source of operating voltage may be connected to the lead w. A circuit will be completed from the lead w through the contact sets of the various relays to one of the fixed contacts of the contact sets associated with the relay 186, depending upon which syllable is stored in the syllable memory. The fixed terminals of the contact sets associated with the relay 186 are connected to the letter decoder 30 (FIG. 1). Operating voltage or ground will then be applied to the appropriate input terminal of the letter decoder. The letter decoder will then operate as described in the above referenced patents and publications to control the automatic typewriter 32 for the printing of the syllable.

It will be assumed, for purposes of illustration, that the syllable "see" was spoken into the microphone 10 (FIG. 1). This syllable is amplified and normalized in the pre-amplifier and amplitude normalizer 12. The frequency spectrum analyzer 14 then splits the signals corresponding to the spoken syllable into a plurality of frequencies channels. The frequency channel components are sorted in the spectral category sorter to provide spectrum category components. These components are switched via the sequence switch 18 into the spectral category memory 20 in which they are stored. The output of this spectral category memory, after the time t₃ is a binary number having 11 bits a to k. In this illustrative case, the bits j and k are binary "zero" bits for the syllable "see," since the syllable growth and decay detector 24 does not detect a rate of syllable growth or decay which is greater than a certain rate. If, instead of the syllable "see" the syllable "by" is sounded, a growth rate greater than the set threshold could be measured by the growth and decay detector

24 and a binary "one" bit for the bit j would be stored in the spectral category memory. Also, at the end of the time t₃, either ground or a source of operating voltage would be applied to the line w in the syllable memory 160 (FIG. 6). The binary number 1 1 0 0 0 1 0 0 0 0 0 or the binary number 1 1 0 0 0 1 1 0 0 0 0 will be stored in the spectral category memory 20 in response to the sounding of the syllable "see." In other words, the relays in the spectral category memory which store the syllable category components corresponding to the binary bits of the last mentioned binary numbers will pull in or not pull in so that their contacts corresponding to either one of the last two mentioned binary numbers.

When a binary "one" bit is stored in the spectral category memory 20, ground is applied to the output terminal of the spectral memory corresponding to that bit. Thus, assuming that the binary number 1 1 0 0 0 1 0 0 0 0 0 were stored in the spectral category memory, ground would be applied to one end of the operating winding for relays 166, 168 and 176. Since the other ends of the operating windings of all the relays 166 to 186 are connected to a source of operating voltage indicated as +B in FIG. 6, relays 166, 168 and 176 will pull in. The other relays 170, 172, 178, 180, 182, 184 and 186 will not pull in. A circuit may then be traced from the lead w through the second sets of contacts of the relays 166 to 186 to the upper fixed contact of the second set of contacts of the relay 166 to 186. Either a source of operating potential or ground will then appear at the output of the syllable memory which corresponds to the syllable "see." The letter decoder 30 (FIG. 1) will then operate to print out the syllable "see" on the typewriter 32 (FIG. 1).

In summary, the syllable memory operates by comparing the binary number stored therein with the binary numbers which are stored in the spectral category memory. If the binary number stored in the spectral category memory satisfies one of the binary numbers stored in the syllable memory (i.e., corresponds thereto) the syllable corresponding to that number stored in the syllable memory is read out into the letter decoder.

From the foregoing description, it will be apparent that there has been provided an improved acoustic apparatus suitable for analysis of speech and other sounds, and for use as a phonetic typewriter. While only one embodiment of the system in accordance with the invention is disclosed herein and its mode of operation explained, variations in the system components, as well as in the system itself, will undoubtedly be apparent to those skilled in the art. Hence, the foregoing should be considered illustrative and not in any limiting sense.

What is claimed is:

1. Acoustic apparatus for encoding into digital form sound which is separable into a plurality of components, said components being sortable into categories of which at least one includes a plurality of said components, which apparatus comprises means for separating said sound into said plurality of components and translating said components into a binary number having a separate binary digit for each of said plurality of components, and means for comparing said binary number with a plurality of defined binary numbers each corresponding to at least one category of said components and deriving an output when said components binary number corresponds to any of said defined binary numbers, said output being a binary digit representing said sound.

2. Acoustic apparatus for encoding in accordance with a digital code sound capable of separation into a plurality of components which are representable by binary digits and which are sortable into selected categories corresponding to defined binary numbers, which categories are representable by binary digits, said apparatus comprising means for separating said sound into said components, and means for sorting said components into said selected categories when the binary digits representing said components pro-

vide binary numbers which correspond to any one of said defined binary numbers, said sorting means providing the binary digits representing said categories into which said components are sorted, said last-named digits representing said sound.

3. Acoustic apparatus for encoding in accordance with a digital code sound capable of separation into a plurality of components which are representable by binary digits and which are sortable into selected categories corresponding to defined binary numbers, which categories are representable by binary digits, said apparatus comprising means for separating said sound into said components, means for sorting said components into said selected categories and for providing binary digits representing said categories when the binary digits representing said components provide binary numbers which correspond to any one of said defined binary numbers, means for storing said category binary digits at discrete intervals during voicing of said sound, and means for obtaining binary number representing said sound from said stored binary numbers.

4. Acoustic apparatus for encoding sound into binary form which comprises means for separating said sound into a plurality of components each represented by a different binary digit, said sound being represented by a binary number including all of said component binary digits, and means for sorting said sound representative binary numbers into different categories each represented by a binary number and each corresponding to a selected group of binary numbers representing certain combinations of said component binary digits, said category binary numbers having a smaller number of digits than said component binary numbers.

5. Acoustic apparatus for encoding into binary form sound which is capable of separation into a plurality of frequency spectrum components, which apparatus comprises means for translating said components into binary numbers each representing a different combination of said components, means for sorting said component combinations into different categories each corresponding to a selected group of said binary numbers, and means for deriving other binary numbers which represent combinations of said categories and having values depending upon whether or not components are sorted into said categories, said other binary numbers having a smaller number of digits than said first-named binary numbers.

6. In apparatus for analyzing sound having means for separating sound into a plurality of sound components, the improvement which comprises a sorter having a plurality of inputs each responsive to a different one of said components and a plurality of outputs corresponding to categories of selected combinations of said components, at least one of said categories including more than one combination of components.

7. In apparatus for analyzing sound having means for separating said sounds into a number of sound components, the improvement which comprises means for deriving a plurality of outputs each corresponding to certain combinations of said components, and means for encoding said outputs into a binary number having a smaller number of digits than the number of said components and representing said sound.

8. Apparatus for sound analysis which comprises means for separating said sound into a plurality of sound components, means for selectively deriving from certain different combinations of said components a plurality of different outputs and for selectively deriving the same output from certain other different combinations of components, and means controlled by said outputs for identifying said sound.

9. The invention as set forth in claim 8 wherein said last-named means includes means for printing a letter representing said sound.

10. Apparatus for sound analysis which comprises means for separating said sound into a plurality of sound components at successive time intervals, means respon-

sive to one or more of the components occurring at each of said time intervals for deriving one of a plurality of outputs, said means being responsive to one or more other components to select said one of said outputs, and means for storing said outputs at each of a plurality of said time intervals and for deriving from said stored outputs a code representing said sound.

11. In apparatus for sound analysis, the combination which comprises means for separating sounds into components, means for deriving those components which exceed a predetermined amplitude level, means for selecting a plurality of combinations of said predetermined level exceeding components which represent sounds having similar characteristics, and means for grouping a plurality of said selected combinations into different categories.

12. Apparatus for recognizing the sounds of speech syllables which comprises means for separating said sounds into a plurality of frequency components each in a different frequency channel, means for sorting said components into groups having similar characteristics, means for grouping a plurality of said sorted groups into categories, and means responsive to said categories for recognizing different syllables from a vocabulary including a plurality of syllables.

13. Apparatus for recognizing the sounds of speech syllables which comprises means for translating said sounds into an electrical signal, a plurality of frequency channels for dividing said signal into a plurality of frequency components, means for translating those components which have signal levels greater than a certain level into a plurality of combinations of signals, means for sorting said combinations of signals into categories each including a plurality of different combinations of said signals, said categories being represented by different electrical signals, and means for deriving from said category signals a code representing said syllables.

14. Apparatus for recognizing the sounds of speech syllables which comprises means for translating said sounds into an electrical signal, a plurality of frequency channels for dividing said signal into a plurality of frequency components, means for translating those components which have signal levels greater than a certain level into further signals, means for sorting said further signals into categories each including a plurality of different combinations of said further signals, said categories being represented by different electrical signals, a memory for storing said category signals, means for applying said category signals to said memory at a plurality of successive time intervals, and means for translating said category signals stored in said memory into a signal representing a syllable stored in said memory.

15. Apparatus for recognizing a sound which comprises means for separating said sound into different frequency components in the spectrum of said sound, means for sorting said frequency components into different spectral combinations of said frequency components and for sorting a plurality of different spectral combinations into a spectral category, means for providing spectral category components representing said categories, and means for encoding said spectral category components into a binary number representing said sound.

16. The invention as set forth in claim 15 wherein said encoding means includes a memory for storing said spectral category components, means for transferring said spectral category components to said memory at each of a plurality of selected time intervals for storage in said memory, and including in addition a syllable memory having storage for a vocabulary of syllables each represented by a different binary code, and means for transferring information in said first-named memory into said syllable memory for storage therein if said information satisfies any of said codes.

17. In apparatus for recognizing the sound of speech syllables, apparatus for obtaining a component which is

present in certain speech syllables and absent in others which comprises means for normalizing said sound in amplitude, and means for providing output signals only when one of the rate of increase and rate of decrease in amplitude of said sound during the voicing of said syllables exceeds a certain predetermined rate and means for adjusting said predetermined rate.

18. In apparatus for recognizing the sound of a speech syllable, apparatus for deriving components representing the growth and decay of said syllable which comprises means for translating said sound into an electrical signal, means for obtaining the envelope of said signal, and means for separately detecting when the growth and decay of said envelope exceed a predetermined rate, said last mentioned means including two channels, one channel detecting only said growth and the other channel detecting only said decay.

19. In apparatus for recognizing the sound of a speech syllable, apparatus for deriving components representing the growth and decay of said syllable which comprises means for translating said sound into electrical signals, means for normalizing the amplitude of said signals, an amplifier having a balanced output for amplifying said signals, a first rectifier polarized to pass signals of positive polarity from said output, a second rectifier polarized to pass signals of negative polarity from said output, first and second differentiating circuits coupled respectively to said first and second rectifiers, a growth channel including said second rectifier and said second differentiating circuit for providing said growth component, and a decay channel including said first rectifier and said first differentiating circuit for providing said decay component.

20. Apparatus for recognizing the sounds of speech

syllables which comprises means for separating said sounds into a plurality of frequency components each in a different frequency channel, means for sorting said components into categories respectively including certain combinations of said components, means responsive to said speech syllable sounds for deriving components of said syllables representing the growth and decay thereof, means for storing at successive time intervals information as to which of said categories include said components, means included in said storing means for also storing said growth and decay components, and means for recognizing different syllables from a vocabulary including a plurality of syllables which correspond to said information in said storing means.

References Cited by the Examiner

UNITED STATES PATENTS

2,181,265	11/39	Dudley	179—1
2,596,199	5/52	Bennett	178—43.5
2,610,295	9/52	Carbrey	178—43.5
2,646,465	7/53	Davis et al.	179—1
2,705,260	3/55	Kalfaian	179—1
2,720,557	10/55	Goodall	178—43.5
3,067,288	12/62	Kalfaian	179—1

OTHER REFERENCES

"Phonetic Typewriter" by Olson et al., IRE Transactions on Audio, July-August 1957, pages 90-95.

ROBERT H. ROSE, *Primary Examiner*.

L. MILLER ANDRUS, WILLIAM C. COOPER,
Examiners.