

[54] VOICE SYNTHESIZER

[76] Inventor: **Richard T. Gagnon**, 307  
Wadsworth, Birmingham, Mich.  
48010

[22] Filed: **July 8, 1974**

[21] Appl. No.: **486,506**

[52] U.S. Cl. .... **179/1 SG; 179/1 SM**

[51] Int. Cl.<sup>2</sup> ..... **G10L 1/08**

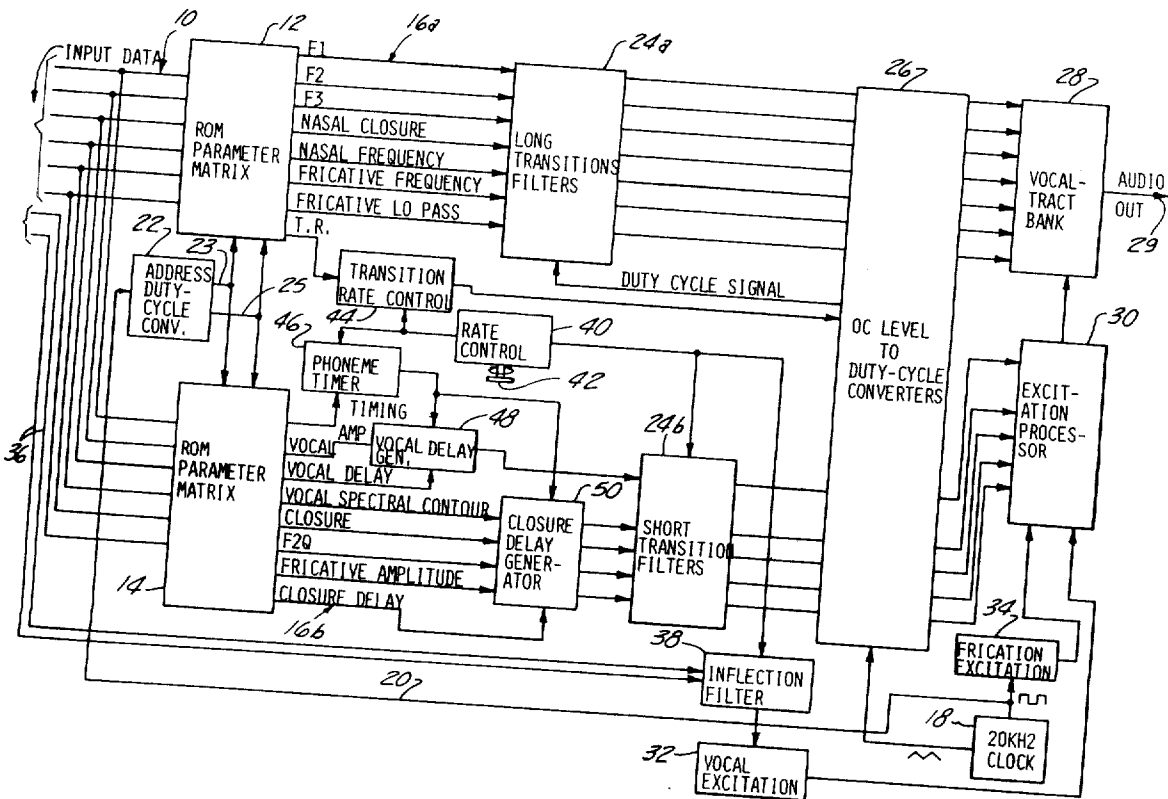
[58] Field of Search ..... **179/1 SA, 1 SG, 1 SM, 15.55 R,  
179/15.55 T**

Primary Examiner—David L. Stewart  
Attorney, Agent, or Firm—Thomas N. Young

[57] **ABSTRACT**

A voice synthesizer of the type set forth in U.S. Pat. No. 3,836,717 wherein the control signals applied to the devices in the vocal tract model take the form of variable pulse width "duty cycle" waveforms. A novel system for producing the duty cycle signals is disclosed. Variable speech rate is provided.

**18 Claims, 6 Drawing Figures**



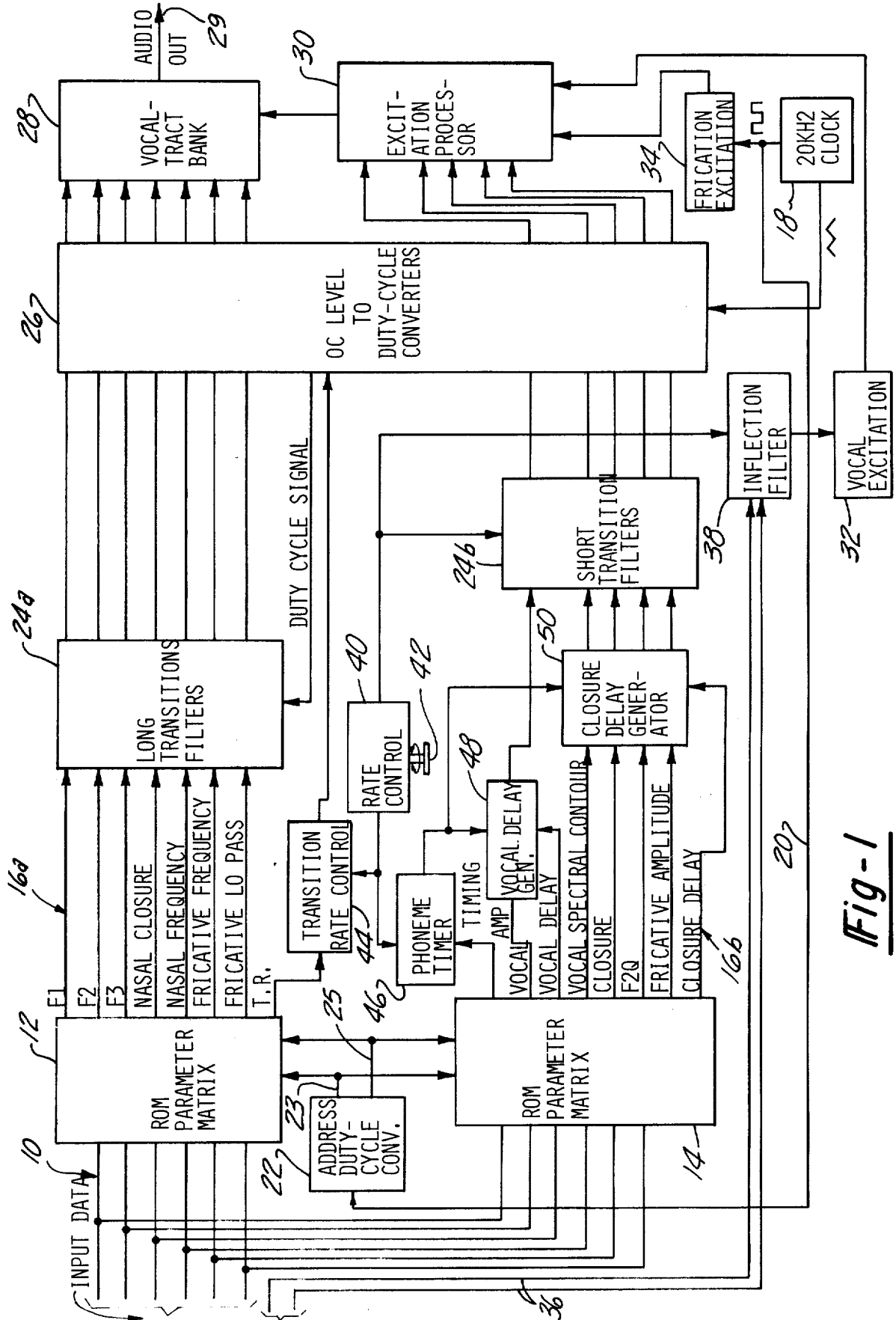
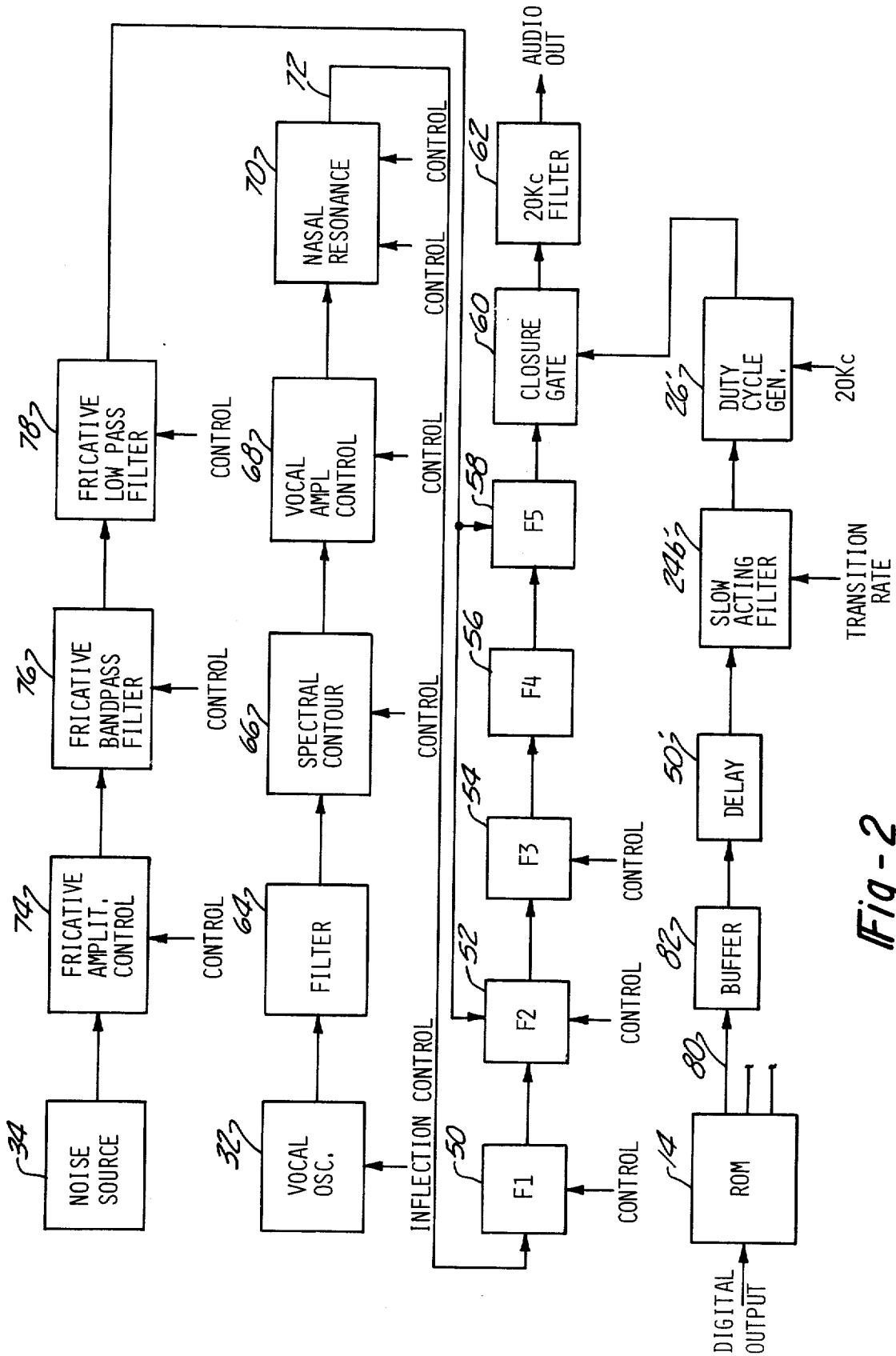
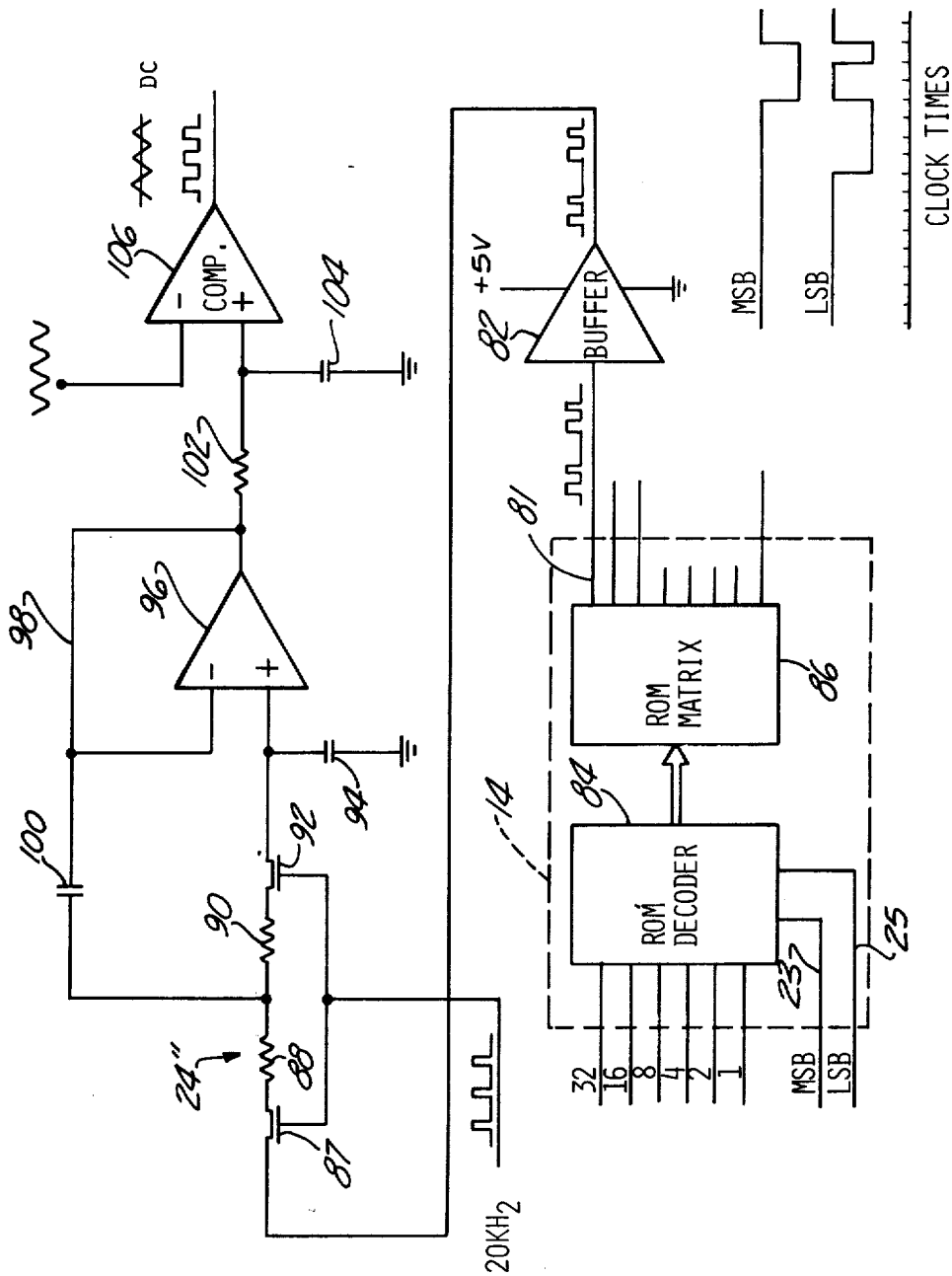


Fig-1



**Fig - 2**



**Fig - 3**

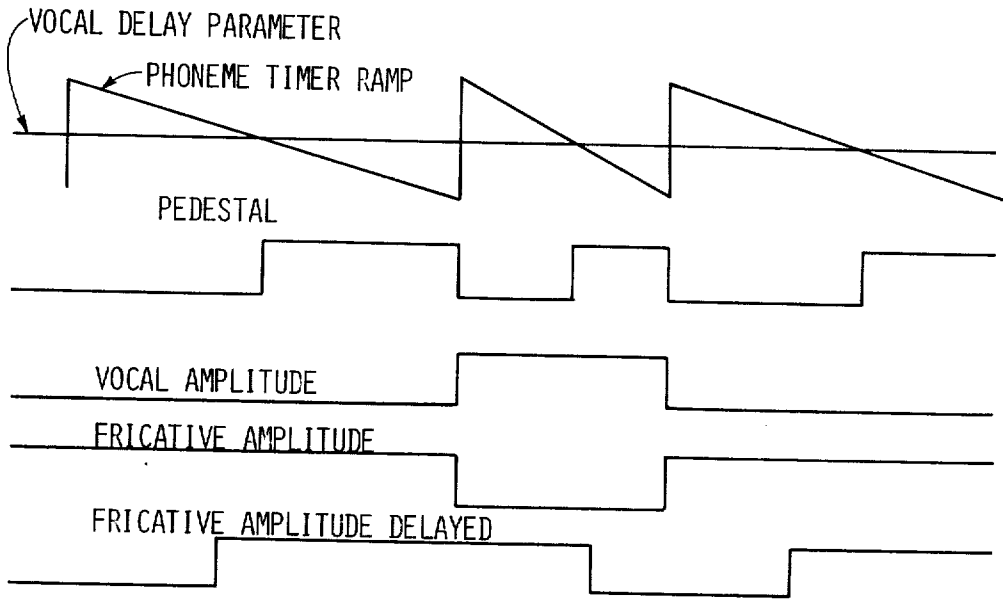


Fig-4

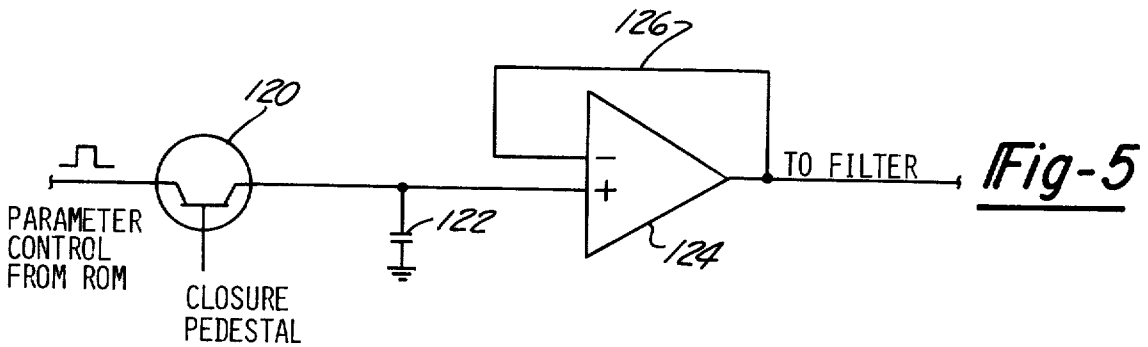


Fig-5

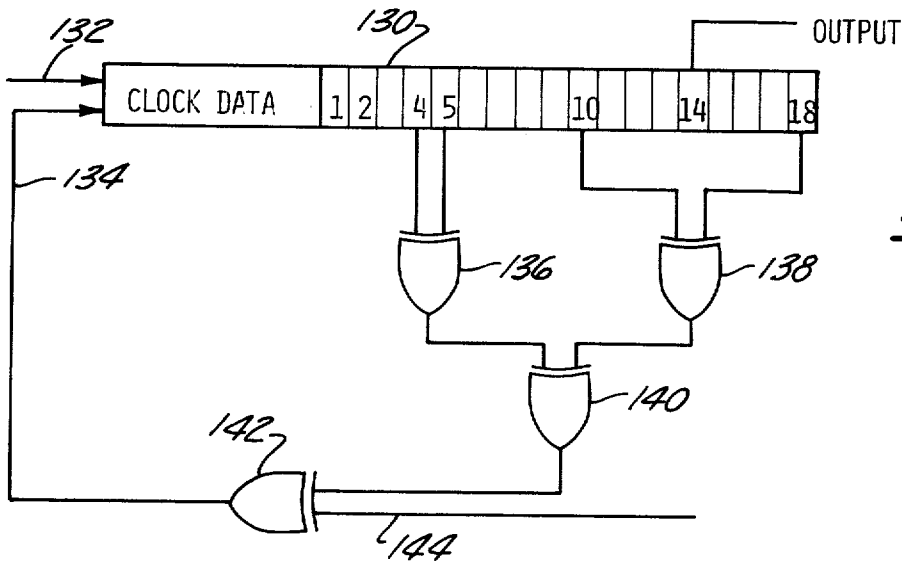


Fig-6

1  
VOICE SYNTHESIZER

INTRODUCTION

This invention relates to voice synthesizers and particularly to improvements in voice synthesizers of the type disclosed in my copending application for U.S. Pat. Ser. No. 274,029, filed July 21, 1972 now U.S. Pat. No. 3,836,717, issued Sept. 17, 1974.

BACKGROUND OF THE INVENTION

The synthesis of human speech by way of readily programmable electronic means presents a vast spectrum of opportunities for application in the field of information transfer of communications. One essential requirement for synthesized speech in any but novelty applications is intelligibility. I have found that one of the critical factors in producing intelligible speech involves the generation of proper dynamic transitions from one phoneme to the next. I have found that in ordinary intelligible human speech the transitions are at least equal in significance to the steady state vocal tract-phonetic conditions since few if any phonemes achieve a steady-state condition for any appreciable time interval. Thus, the quality of the transitions contributes to intelligence as well as realism.

Other prior art practitioners have also apparently recognized the importance of interphoneme transitions. In one prior art system, elaborate electronic means are provided for producing piecewise linear approximations of the phoneme transition waveforms of synthesized speech. At least one other system involves a highly complex electronics system for making a running comparison between adjacent phonemes so that specific and predefined phoneme interaction waveforms may be generated in response to coded representations of the various phoneme sequences which occur in a given speech pattern.

I have found that it is not necessary to resort to elaborate and complex electronics to produce satisfactory interphoneme transitions. I have greatly simplified the prior art systems which, at the same time, I have produced superior interphoneme transitions by generating a succession of analog control signals defining steady-state phoneme parameters and by passing these signals through slow-acting filters which prevent the steady-state values from being achieved. As is more fully set forth in my copending application for patent Ser. No. 274,029, I employ a vocal tract model comprising tunable resonant filters and amplitude control devices responsive to analog control signals for defining the various constituents of each phoneme in speech pattern. I generate the control signals first in digital form and I convert the signals to dc voltage levels in my illustrated embodiment by means of digital to analog converters. Thereafter, I pass the analog outputs through relatively slow acting filters; i.e., filters whose transfer characteristics would, in sufficient time, duplicate the end parameters of the control signal but which, within the ordinary phoneme duration intervals, will prevent the analog control signals as applied to the vocal tract model from reaching the steady-state or target values. In this fashion, the sluggish acting filters tend to produce smooth, continuous glides between phonemes in much the same manner as the human vocal tract produces smooth transitions between phonemes in spoken speech.

2

BRIEF SUMMARY OF THE INVENTION

I have now found that it is possible to further improve upon the electronics system in speech synthesizers of the type set forth in my copending patent application Ser. No. 274,029 particularly with respect to the generation of the analog control signals from digital input quantities. In general, I accomplish this by providing means responsive to digital phoneme parameter specification signals for generating variable duty cycle wave forms, the average values of which are the analog equivalent of the digital signals. In my preferred embodiment hereinafter described in detail, the variable duty cycle signals are generated by serializing digital signal quantities using a binary progression of weighted time intervals. Thereafter, I filter the serialized variable duty cycle waveforms to produce an analog signal for application to the analog control devices in the vocal tract model.

I have also found that it is possible to provide the facility for variable speech rates in a synthesizer of the type set forth in my copending application for patent Ser. No. 274,029 while at the same time preserving all of the advantages of the simple filters for producing the interphoneme transition. In general, I accomplish this by variable timing means for varying the phoneme intervals and thereby the speech rate and further by tuning the slow acting filters such that the response times track with the varying speech rates; i.e., the response times are made shorter for higher speech rates and proportionately shorter phoneme intervals, but the ratio of phoneme interval to response time remains about the same for a given phoneme. In the specific embodiment hereinafter described I accomplish this through the use of analog gates in series with the resistive components of the slow acting filters and by applying a variable duty cycle, high-frequency chopping signal to the analog gate thereby simulating a varying electrical parameter; i.e., resistance. In this fashion, I avoid the problem which might otherwise occur in my prior system if the actual speech rate were set either too low or too high in relation to the filter response time.

I have also discovered other improvements which might be made to my prior voice synthesizer including the use of simple analog gates in the vocal tract model so as to permit control by means of a smoothly varying duty cycle signal, and the delay of certain parameters such as voice constituent amplitude, fricative closures and certain other parameters thereby to produce still more realistic speech. I also disclose herein the use of a highly simplified noise source for generating unvoiced phoneme constituents. These improvements may, of course, be employed in various combinations with or without the other aspects of this invention as described herein. The various features and advantages of the invention will be best understood from a reading of the following specification.

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 is a block diagram of a voice synthesizer system employing the present invention;

FIG. 2 is a more detailed block diagram of a portion of the system showing the specific element of the vocal tract model;

FIG. 3 is a partial circuit diagram of representative portion of a single control signal generating channel;

FIG. 4 is a timing diagram of wave forms relating to the circuits of FIGS. 1 through 3;

FIG. 5 is a circuit diagram of a delay system for producing the effects shown in FIG. 4; and

FIG. 6 is a circuit diagram of a noise source.

#### DETAILED DESCRIPTION OF THE SPECIFIC EMBODIMENT

Looking now to FIG. 1, the major portion of a system block diagram is represented. This system, like the system of my copending application Ser. No. 274,029 may be operated by using any of several types of input means including business machines, computer, or phonetic key boards capable of generating a sequence of digital signals representing the various phonemes to be selected thereby to make up a given speech pattern. Since such input means are described in the prior art as well as in my copending application, I have omitted a specific description from this text. The phoneme selection signals are transmitted by way of the six input lines 10 to solid state read-only memories 12 and 14, each having six input address lines and eight output lines. It will be understood that I use two separate memories only because my specific input/output requirements were not met by a currently available device, there is no technical reason why a single unit would not suffice if available. The memories 12 and 14 respond to the phoneme selection addresses to generate phoneme parameter control signals of an analog character on the output lines 16. As hereinafter described in greater detail, the signals appearing on output lines 16 are sequences of serialized digital signal quantities wherein the various bits in the repeating series are time-weighted according to a binary progression. Thus, the output signals have an average value which is the analog equivalent of the digital input quantities addressing those outputs. Output lines 16a individually control eight phoneme parameters including output spectral frequencies, input forcing function frequencies, nasal closure, and transition rate. The eight output lines 16b control eight more phoneme parameters including timing, amplitude, delay, spectral contour, closure, and band width. A total of 16 phoneme parameters, each capable of 16 different values, are employed in the system illustrated herein to control the phonetic output.

The analog control quantities on output lines 16a and 16b are connected through the filters 24a and 24b, respectively, the response times of filters 24a being tuned so as to be long relative to the typical phoneme intervals which are selected for a given speech rate. Filters 24b also produce a damped response to step inputs but to a lesser degree than filters 24a. The nominal setting of the filters 24a and 24b is such that the response times thereof bear some predetermined relationship to the phoneme times peculiar to a given speech rate, the result being that the frequency parameters output from ROM 12 on lines 16a are seldom, if ever, realized during the phoneme interval to which they directly relate. Similarly, the step inputs to filters 24b are smoothed by the filter response characteristics. Therefore, the filters 24 produce the phoneme control signal smoothing which is described in my copending application for all speech rates, again, I emphasize that this is most important for the vocal tract frequency control signals F1, F2, and F3. The major difference here lies in the fact that these filters 24 are tunable for varying speech

rates. As in the system of my copending application, the transfer characteristics of filters 24 are such as to eventually produce an output which approximates the input, given enough time to respond; i.e., the level eventually reaches a steady state amplitude level related to the amplitude of the signal at the input.

The seven output signals from filters 24a and the five output signals from filters 24b are connected through the duty cycle converters 26 which convert the smooth slowly varying analog dc levels into fixed-frequency pulse trains wherein the duration or width of the pulses varies according to the input dc levels. This duty cycle or "pulse width modulated" signals are then applied to the various devices in the vocal tract model 28 to produce the audio speech output on line 29. The lower five parameter control signals relating to amplitude, spectral contour, fricative frequency and spectral shape are applied to the filter bank in the vocal tract 28 through an excitation processor 30 comprising analog control devices for the control of the quantities indicated.

The two basic constituents of synthesized speech are the vocal and fricative forcing functions which are provided sources 32 and 34, respectively. Source 32 is a source of audio signals which are used to produce the voiced phoneme constituents. Source 32 is connected to the vocal tract filter bank 28 through the excitation processor 30. The fricative excitation source 34 is a noise source hereinafter described in greater detail and is also connected to the vocal tract filter bank 28 through the excitation processor 30.

The duty cycle outputs on lines 16 are generated by means of basic timing apparatus including a 20 KHz clock source 18 having output line 20 connected to the duty cycle conversion unit 22 and having output lines 23 and 25 connected to the read-only memories 12 and 14. As will be hereinafter described in greater detail, the signals on lines 23 and 25 form part of the phoneme addresses and operate to serialize four selected bits of stored data onto each of the parameter output lines 16 in a binary progression wherein the first bit is assigned eight clock times, the second bit is assigned four clock times, the third bit is assigned two clock times, and the last and final bit is assigned a single clock time. More details on the manner in which this specific conversion is accomplished will be described with reference to FIG. 3.

Inflection signals are input from the programming means by way of lines 36 and connected to the inflection filter 38, the output of which is connected to the vocal excitation source 32 to vary the frequency or pitch of the vocal excitation source output. This produces inflection variations in the audio output on line 29.

As previously described, a feature of the present system is the ability to produce speech at varying rates from relatively fast to relatively slow rates without loss of intelligibility. The basic rate control signal is controlled by unit 40 having a manual tuning dial 42. The speech rate signal is a duty cycle signal; i.e., a time varying wave form of fixed frequency (20 KHz) but variable pulse width and is connected to the slow acting filter bank 24b, the inflection filter 38 and a transition rate control unit 44 and a phoneme timer unit 46 which produces an output ramp varying from 5 volts to 0 volts in a period or interval which varies with phoneme interval timing. It will be noted that the relative timing parameter from read-only memory 14 is one of the 16 pa-

rameters selected by the stored data in the memories and is applied to the phoneme time 46 to establish the slope of the output ramp from the timer 46. Thus, phoneme timing varies not only with speech rate on an across-the-board basis, but also from phoneme-to-phoneme at a give speech rate. The ramp output from timer 46 is connected to a vocal delay generator 48 which functions to delay the vocal amplitude control parameter and to a closure delay generator 50 which controls a time delay of the vocal spectral contour, closure, bandwidth, and fricative amplitude control parameters, all of which are independent of the vocal tract resonant functions controlled by the memory 12. The output rate of control unit 40 is also connected to the transition rate control so that the transition rates are taken into account in controlling the response times of the filters 24a whereby those response times track with the phoneme intervals.

Looking now the block diagram of FIG. 2, the details of the vocal tract model 28 will be shown with greater specificity. The vocal tract filter bank comprises cascaded resonant filters 50, 52, 54, 56, and 58 which produce the frequency poles in the output spectrum of any given phoneme. Each of the filters may be implemented in the form of a two pole filter as is more fully set forth in my copending application Ser. No. 274,029. It will be noted, however, that in the present embodiment of my invention the filters are connected in series rather than in parallel as in my previous embodiment. I find that this produces a certain advantage with respect to energy distribution between the poles and creates a more realistic speech output. It is to be understood, however, that I do not intend the present invention to be limited to any particular arrangement of resonant filters but rather that I presently find the cascaded arrangement of filters to produce superior results. Filters 50, 52, and 54 are tunable so as to vary the positions of the poles in the output phoneme spectrum, whereas filters 56 and 58 are fixed pole filters. The cascaded arrangement of filters is connected through a closure gate 60 which is subject to a control signal, and a 20 KHz filter 62 which filters out the control signal carrier which might otherwise appear in the output waveform. Again the output signal appears on line 29 and corresponds with FIG. 1.

The entire vocal tract model, of course, comprises the vocal oscillator 32, the noise source 34 and the respective control channels for the forcing functions. Vocal oscillator 32 is connected through a filter 64 and a spectral contour filter 66 which is subject to tuning by an externally derived control; i.e. the "vocal spectral contour" control signal produced on the fourth output line of read-only memory 14. The vocal oscillator signal is also connected to a vocal constituent amplitude control unit 68 which is an analog gate as hereinafter described in greater detail. The unit 68 is also subject to the externally derived control signal; i.e., the second output signal of read-only memory 14. Finally, the signal is passed through a nasal resonance filter 70 which is subject to two control signals, the "nasal closure" and the "nasal frequency" signals which are derived on the fourth and fifth output lines of read-only memory 12. The output of the nasal resonance filter is connected by way of line 72 to the input of the vocal tract filter bank; i.e., at the input of filter 50 as shown in FIG. 2.

The fricative noise source 34 is connected through the fricative amplitude control device 74 which is subject to external control, the fricative band pass filter 76 which is subject to external control, and the fricative low pass filter 78 which is also subject to external control. The amplitude controlled and filtered fricative forcing function is injected into both the F2 and F5 filters 52 and 58 in the vocal tract filter bank as shown.

The block diagram of FIG. 2 also contains a portion of a representative control signal generating channel, in this case the channel which generates the closure control signal applied to gate 60 in the vocal tract filter bank. The control signal channel comprises the read-only memory 14 which receives the digital input signal and which produces analog output signals on the various output lines thereof. The output line 80 of interest is connected through a buffer amplifier 82 to establish precision voltage limits on the duty cycle signal and is thereafter applied to the closure delay generator unit 50' as previously described. The output is applied through the slow acting tunable filter 24b' and thereafter through the duty cycle converter 26'. From there the duty cycle signal is applied directly to the closure gate 60 for control over the closure function. It is to be understood that the term "duty cycle signal" as used therein refers to a fixed frequency pulse train which varies between two relatively fixed amplitude levels with varying pulse widths.

Looking now to FIG. 3, another representative control signal channel will be described in still greater detail. In FIG. 3, the read-only memory unit 14 is shown divided into decoder and output matrix sections 84 and 86, respectively. The decoder unit 84 receives the phoneme address on the six input address lines having the binary weighted address values shown in the drawings. To select phoneme number 13, high input signal values are applied to the "eight", "four", and "one" input signal lines while low signal values are applied to the remaining lines. Other addresses are similarly selected. The input signal polarity convention may, of course, be reversed depending upon the specific circuitry employed. In addition, the timing signals on lines 23 and 25 are applied to the decoder 84 as further address constituents, the "MSB" signal on line 23 having the timing characteristic illustrated, and the "LSB" signal on line 25 having the modified timing characteristic also illustrated. More specifically, both the MSB and LSB signals have 15 clock time periods broken into eight, four, two, and one clock time segments. The MSB signal has a high value for the first two segments and a low value for the last two segments whereas the LSB signal has a high value for the first and third segments and a low value for the second and fourth segments. These nonvarying timing signals are applied to the decoder section 84 of memory 14 for all input signal combinations and operate to complete the address inputs to the decoder 84.

Specifically, the combination of signals applied to the six binary weighted input lines select groups of four output bits for each of the eight output lines from the matrix section 84 of read-only memory 14. The time distribution or order of the four selected bits for each parameter is controlled by the time varying relationship between the MSB and LSB signals and functions to distribute the bits in an eight-four-two-one clock time sequence on each of the parameter output lines of which line 81 is the selected example. In other words, the first



bit selected by the six bit address appears for eight clock times, the second bit selected appears for four clock times, the third bit selected appears for two clock times, and the last bit selected appears for one clock time. This has the effect of producing a serialized and binary-weighted duty cycle signal on the output line 81, the average value of which varies between 0 and 15 as increments of the maximum output voltage values; i.e., 5 volts. Thus, each analog signal value is spaced from the adjacent analog signal values by approximately one-third volt in amplitude, an easily detected amplitude variation for control purposes.

Output line 81 is connected to the input of the buffer amplifier 82 which, as shown in FIG. 3, has the upper limit input pin connected to a precision 5 volt source and the lower limit input pin connected to ground. Thus, the duty cycle signal which is input to the amplifier 82 is reproduced at the output but between precisely defined voltage limits of 5 volts and 0 volts so as to insure accuracy in the average value of the duty cycle signal. The advantages of the duty cycle signal conversion which is employed in the present embodiment of the invention are substantial in that it results in the generation of four bits for each phoneme parameter in series yet at the same time requires only two read-only memory units to generate all 64 bits. Moreover, the particular serialized generation of the 16 four-bit groups requires no latch devices to hold the four bits for simultaneous application to a digital-to-analog converter. In other words, the approach of the present invention eliminates the need for latches as well as separate digital-to-analog conversion devices such as resistor ladder networks. Of course, an alternative approach would be to employ a sufficient number of read-only memory units to generate all 64 bits at once in parallel but the economic as well as spatial requirements of this approach limit practicality as will be apparent to those skilled in the art.

The duty cycle signal comprising the binary weighted distribution of four bits on the parameter output line 81 is connected from the buffer amplifier 2 to a tunable slow-acting filter 24'' which forms part of one of the filter banks either 24a or 24b. A delay device may be employed in the connection. The filter comprises an analog gate 87 having the primary terminals connected to pass the duty cycle signal therebetween, a resistor 88, a second resistor 90, a second analog gate 92, and a shunt capacitor 94 connected to the positive input of an operational amplifier 96. The output of the amplifier 96 is connected through feedback path 98 back to the negative input of the amplifier and also through a capacitor 100 to the junction between the resistors 88 and 90. The transition rate signal from unit 40, chopped at the 20 KHz rate is applied to the control terminals of the two analog gates 87 and 92. The control signal applied to the analog gates 87 and 92 operates to render the gate conductive and non-conductive at a very high frequency relative to the highest frequency component in the input duty cycle signal and varies in its own duty cycle in accordance with the desired transition rate. Thus, the average on-off time ratio of the gates 86 and 90 is varied in direct proportion to the transition rate signal. This has the effect of varying the apparent resistance of the tunable slow-acting filter 24'' in accordance with the transition rate signal so that the response time of the filter tracks with the desired speech rate; i.e., it will be recalled from the description

of FIG. 1 that the setting of the rate control unit 40 effects the transition rate control unit 44 which in turn controls the duty cycle or pulse width of the chopped signal applied to the control terminals of the gates 86 and 92.

Accordingly, the output of the amplifier 96 is a dc voltage the amplitude of which varies with the average value of the duty cycle signal which is input to the filter 24''. Of course, the signal input to filter 24'' is changing and thus the output dc level is substantially continuously varying as well. The output of amplifier 96 is preferably connected through a "glitch" filter comprising a series resistor 102 and a shunt capacitor 104 to get rid of spurious signals. The output of the glitch filter is connected to the comparator amplifier 106 which forms part of the unit 26 illustrated in FIG. 1. This unit comprises a comparator amplifier having the positive input pin connected to receive the varying dc voltage level and the negative pin connected to a 20 KHz sawtooth voltage wave which varies between 0 and 5 volts. Again, it will be observed that the 20 KHz signal operates to phase synchronize all duty cycle signals in the system. The output of comparator amplifier 106 is a fixed frequency pulse train wherein the pulse widths vary in accordance with the portion of the 20 KHz sawtooth which exceeds the dc voltage level applied to the positive input terminal of the comparator amplifier 106. This is a function of the amplitude of the dc signal. Such means to convert from dc levels to duty cycle signals are well known to those skilled in the art and will not be described in greater detail herein.

It is to be understood that the duty cycle signal output from ROM 14 is converted to a dc voltage level and then back to a duty cycle signal only to facilitate the filtering function at unit 24. If satisfactory filtering can be accomplished by operating on the duty cycle signal directly, such conversion may be eliminated.

The reconverted duty cycle signal which represents the phoneme control signal is then applied directly to the device in the vocal tract model which is to be controlled to produce the desired contribution to the particular selected phoneme. The signal might, for example, be the F1 signal which is applied to control the position of the frequency pole of the filter 50; it could be any one of the other 15 control signal quantities which are generated except, of course, the transition rate signal, the timing signal, the vocal delay signal, or the closure delay signal, none of which are connected directly through tunable slow-acting filters. The advantages of the reconversion from dc to duty cycle is to eliminate the requirement for the complex analog multipliers which are disclosed in my copending patent application and to permit the use of simple analog gates to perform proportional control and effective signal multiplication on a duty cycle basis. The economies in terms of cost, complexities and spatial requirements of this approach will be apparent to those skilled in the electronics arts.

Another important feature of my voice synthesizer as disclosed herein is the delay of excitation functions so as to match the filtering events with respect to the timed relationship between adjacent phonemes or phonemes constituents of different types. For example, in a transition from vocal to fricative phonetic constituents such as one finds in the pronunciation of the letter "s" it is desirable to delay the excitation of the fricative portion so as to be spaced from the amplitude decaying vocal portion of the letter. The same may be true in re-

verse; i.e., in fricative to vocal transitions as well. FIG. 4 discloses the saw-tooth shaped phonetic timer ramp signal which is output from the phonetic timer 46 in the circuit of FIG. 1 and which controls phoneme duration as previously described. When compared to a dc voltage level representing the vocal delay command from ROM 14, it can be seen from the second line of FIG. 4 that a "pedestal" can be generated. This pedestal signal is used in the circuit of FIG. 5 as hereinafter described.

FIG. 4 also shows the typical relationship between a fricative to vocal transition wherein the amplitude of the fricative forcing function is shown to drop sharply at exactly the same time as the amplitude of the vocal forcing function rises. The reverse is true at the end of the second phoneme time interval. To utilize these excitation functions without modification would produce unrealistic speech as the forcing functions would not correlate properly in time with the control parameters which are derived through the slow acting filters as previously described. Accordingly, the fricative forcing function is delayed such that the amplitude rise occurs later in the phoneme time as shown in FIG. 4.

To accomplish the delay function, the closure delay generator 50 and the vocal delay generator 48 in FIG. 1 are employed. Since these are similar, if not identical, in implementation, FIGS. 4 and 5 will be described as representative of either or both. In FIG. 5 the forcing function or control parameter from the ROM is shown applied to the input of a switch 120 while the closure delay pedestal derived from the comparison between the phoneme timer ramp and the closure delay signal as previously described with reference to FIG. 4 is applied to the control terminal of the switch 120. The portion of the forcing function, e.g., fricative amplitude, which is passed during the on time of the switch 120 is stored in a capacitor 122 and applied to the positive input of an operational amplifier 124 having feedback path 126. The output of the operational amplifier is a signal related to the input control parameter but delayed by a sufficient interval as to properly synchronize the excitation events with the filtering events in the output of the synthesizer. It will be understood that a variety of approaches to this delay function can be employed and that the delay concept in general is one of several techniques described herein which may be applied in various combinations for the purpose of accomplishing more realistic speech in the synthesizer process. The vocal forcing function is similarly delayed.

FIG. 6 is a schematic circuit diagram of an improved noise source 34 as shown in FIGS. 1 and 2. The circuit of FIG. 6 is capable of generating a pseudo-random mixture of frequencies having excellent spectrum and amplitude characteristics for use in voice synthesis.

The circuit comprises an 18 bit shift register 130 having a 20 KHz "clock" input 134. Taps nos. 4 and 5 are connected to opposite inputs of exclusive-OR gate 136 while taps 10 and 18 are connected to the inputs of similar gate 138. The outputs of gates 136 and 138 are OR'd through gate 140. The output of gate 140 is exclusively OR'd through gate 142 with the 1.33 KHz input waveform on line 144. The register output may be taken from any tap, and is shown taken from tap no. 14. The output is a pseudo-random function, the period of recurrence for which is long enough to look like purely random noise. The use of the 1.33 KHz signal is effective to avoid a "lockup" condition wherein all of

the gate outputs and register inputs are the same and the sequence fails to progress.

It is to be understood that the invention has been described with reference to specific aspects of a specific embodiment and accordingly the foregoing description is not to be construed in a limiting sense.

What is claimed is:

1. In a voice synthesizer for phonetically synthesizing human speech:

10 a vocal tract model comprising tunable filters and amplitude control devices responsive to input signal quantities to determine the frequency and amplitude parameters of phonemes to be output therefrom;

15 input means for specifying selected phonemes to be output from said vocal tract model and said parameters thereof in a predetermined order;

and control signal forming means connected between said input means and said tunable filters and amplitude control devices of the vocal tract model for producing a plurality of variable duty cycle waveforms as input signals for controlling the frequency and amplitude characteristics of the selected phonemes.

25 2. Apparatus as defined in claim 1 wherein said input means comprises means for specifying a phoneme code having a plurality of bits, said signal forming means between the input means and vocal tract model including means for cyclically outputting said bits in series according to a predetermined timing sequence having a numerically progressive order whereby said bits define said duty cycle waveforms, the average value of which varies according to the selected bits of said code.

30 3. Apparatus as defined in claim 1 wherein said signal forming means comprises filter means for producing a dc signal related to the average value of the duty cycle waveforms.

4. Apparatus as defined in claim 3 further including means for converting said dc signal to a second variable duty cycle signal of fixed frequency and amplitude for application to said vocal tract model.

5. Apparatus as defined in claim 1 including buffer means connected in said signal forming means for establishing precision voltage levels for said variable duty cycle signals.

6. Apparatus as defined in claim 1 wherein said input means includes timing control means for selecting a relative time interval for each specified phoneme, and timing control means for varying the overall rate of phonetic production while preserving the relative timing between said phonemes.

7. Apparatus as defined in claim 6 wherein said signal forming means comprises filter means for producing a dc signal related to the average value of the duty cycle waveforms.

8. Apparatus as defined in claim 7 wherein said filter has a response time which is slow relative to the phoneme repetition rate.

9. Apparatus as defined in claim 8 wherein said timing control means includes means for varying the response time of said filter to track with said overall rate of phonetic production.

10. Apparatus as defined in claim 1 wherein said signal forming means comprises at least one storage facility having input address means and output control means some but not all of the input means being connected to receive phoneme address signals, said facility

11

12

having stored therein a plurality of phoneme control parameters which are applied to said output means according to the address signals, each phoneme address signal combination applied to said some of the input means being effective to select a plurality of parameter control signals for application to an output, timing signal means connected to the remaining input means for applying thereto timing signals for applying the selected control signals to the output individually and in a sequence having a binary-weighted time order such that the average value of the total sequence is a function of the particular address signal combination.

11. Apparatus as defined in claim 10 wherein the sequence timing is in the order 8-4-2-1 such that the average value may occur in the range of 0 to 15.

12. Apparatus as defined in claim 10 wherein the control signals on said output comprise an electrical wave-form which varies between two relatively fixed amplitude levels.

13. Apparatus as defined in claim 12 further including buffer means connected to receive the control signal sequence for reproducing said sequence but with a close-tolerance on said amplitude levels.

14. Apparatus as defined in claim 13 including a filter connected to receive the control signal sequence and to output a dc signal related to the average value thereof, the filter having a response time which is slow

relative to the rate at which the average value varies from one selected value to another.

15. Apparatus as defined in claim 14 further including timing means for controlling the rate of phoneme selection, said means being connected to the filter for varying said response time to track with the rate selected.

16. Apparatus as defined in claim 15 wherein the filter comprises gate means connected to transmit the parameter signal value, and a control terminal connected to receive a high frequency signal having a variable duty cycle thereby to control the ON-time of the gate means.

17. Apparatus as defined in claim 16 further including means connected to said filter for converting the dc signal to a variable duty cycle signal for application to the vocal model.

18. Apparatus as defined in claim 1 including as part of said vocal tract model a source of voiced phonetic excitation quantity and a source of unvoiced phonetic excitation quantity, and means responsive to the addressing of a consecutive phonetic sequence of voiced and unvoiced phonetic quantities for producing a predetermined delay in the excitation of at least one of said quantities.

\* \* \* \* \*

30

35

40

45

50

55

60

65