

May 12, 1970

J. L. FLANAGAN

3,511,932

SELF-OSCILLATING VOCAL TRACT EXCITATION SOURCE

Filed Aug. 29, 1967

3 Sheets-Sheet 1

FIG. 1

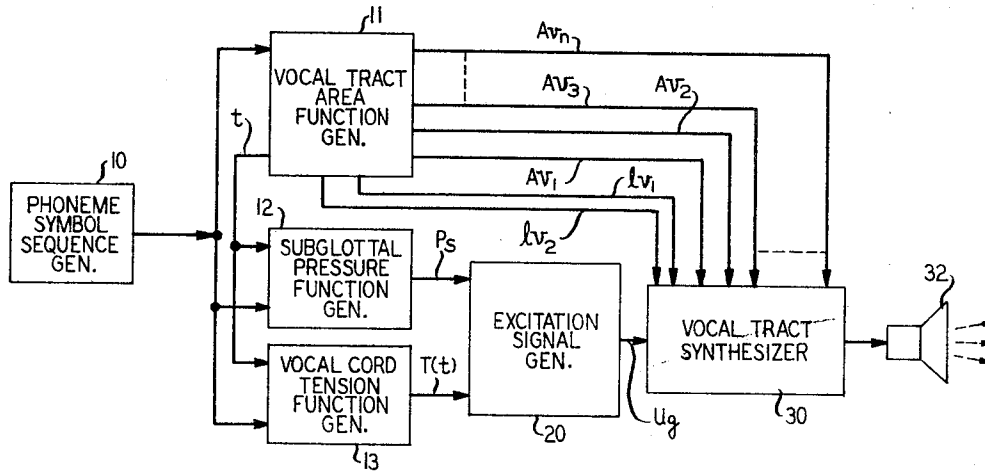


FIG. 2

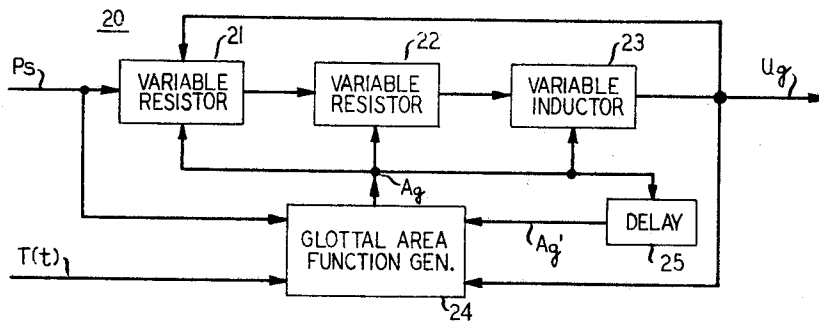


FIG. 3A

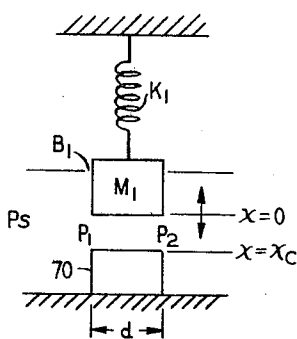


FIG. 3B

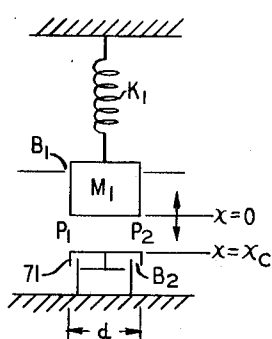
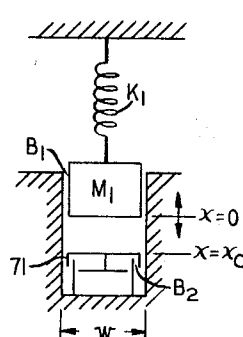


FIG. 3C



INVENTOR
J. L. FLANAGAN
BY *G. E. Ulrich Jr.*
ATTORNEY

May 12, 1970

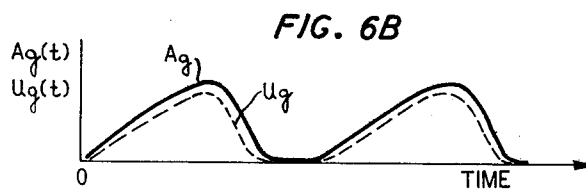
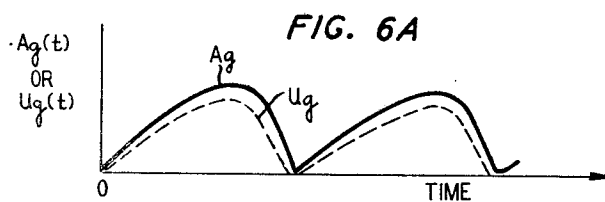
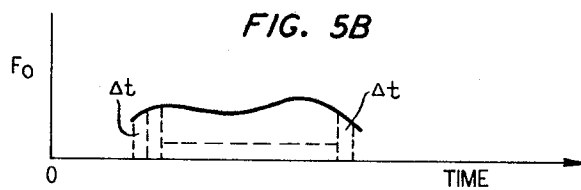
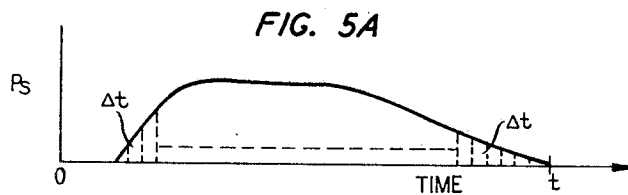
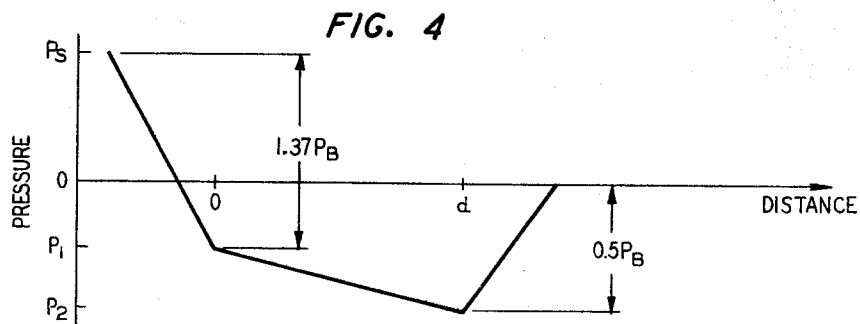
J. L. FLANAGAN

3,511,932

SELF-OSCILLATING VOCAL TRACT EXCITATION SOURCE

Filed Aug. 29, 1967

3 Sheets-Sheet 2



May 12, 1970

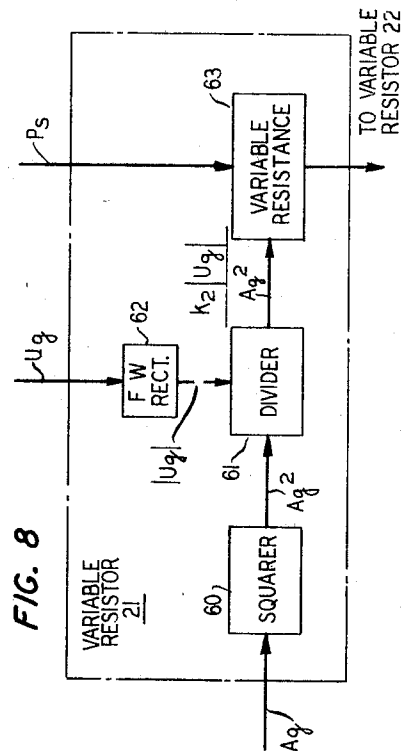
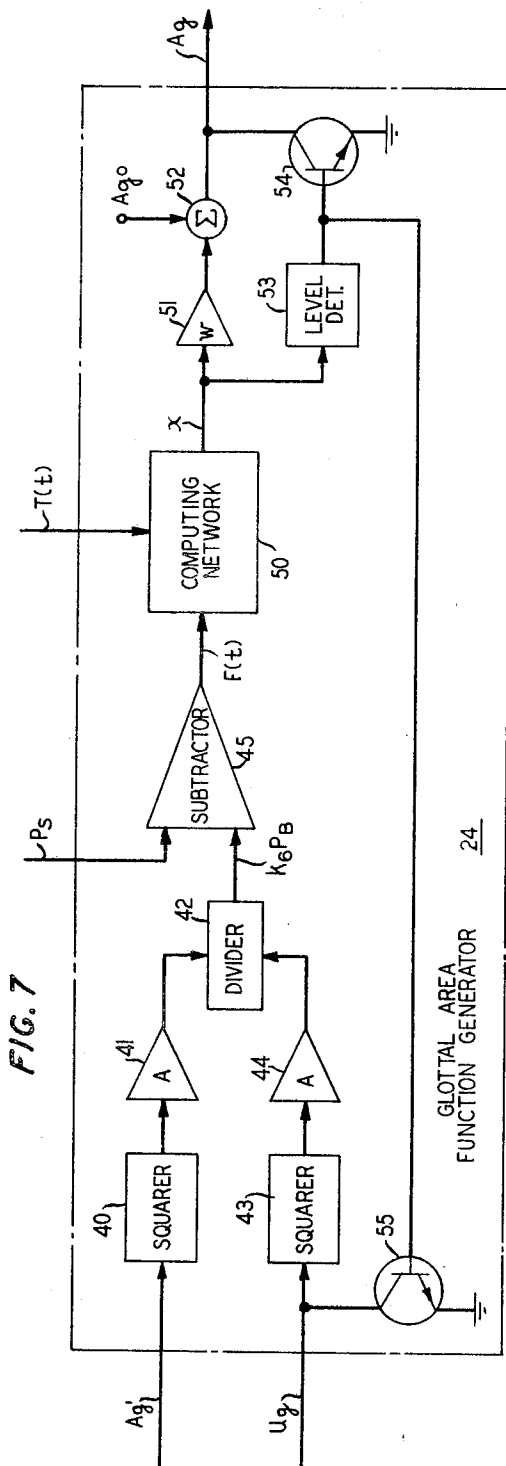
J. L. FLANAGAN

3,511,932

SELF-OSCILLATING VOCAL TRACT EXCITATION SOURCE

Filed Aug. 29, 1967

3 Sheets-Sheet 3



1

3,511,932 SELF-OSCILLATING VOCAL TRACT EXCITATION SOURCE

James L. Flanagan, Warren Township, Somerset County, N.J., assignor to Bell Telephone Laboratories, Incorporated, Murray Hill and Berkeley Heights, N.J., a corporation of New York

Filed Aug. 29, 1967, Ser. No. 664,130

Int. Cl. G10l 1/00

U.S. Cl. 179—1

10 Claims

ABSTRACT OF THE DISCLOSURE

Speech synthesizers often produce unnatural sounding speech. This is due in large measure to the problem of developing a suitable vocal tract excitation signal. This problem is overcome by representing the human vocal cords as a second order oscillatory system that is driven by a nonlinear forcing function of subglottal pressure. Accordingly, apparatus is utilized that generates an excitation signal in response to the physiological parameters of subglottal pressure, vocal cord tension, and vocal tract configuration.

BACKGROUND OF THE INVENTION

Field of the invention

This invention relates to apparatus for synthesizing speech and, more specifically, to apparatus for generating a vocal tract excitation signal.

Description of the prior art

Problems arise in producing synthetic speech of acceptable quality due to difficulties in duplicating human vocal tract excitation information. Several systems have been proposed which attempt to produce natural-sounding speech through synthesis. For example, a system which reproduces prerecorded phonetic sequences to form word groups in response to a coded input signal is disclosed in U.S. Pat. 2,771,509, issued to H. W. Dudley and C. M. Harris on Nov. 20, 1956. An improvement of the Dudley-Harris apparatus is disclosed in U.S. Pat. 3,158,685, issued to L. J. Gerstman and J. L. Kelly, Jr. on Nov. 24, 1964. In the Gerstman-Kelly apparatus, phonetic symbols are converted to electrical signals which control a formant synthesizer to reproduce speech. Synthesized speech produced by these systems, however, does not attain a quality comparable to human speech; it is generally characterized as "computer talk" or "machine talk." Thus, although speech so generated may be acceptable functionally, it is unsatisfactory for general application due to its unnatural sound.

Still another suggested approach to improving the quality of synthesized speech is disclosed in chapter 3 of Speech Analysis, Synthesis and Perception, by J. L. Flanagan, Academic Press, 1965. There, in an attempt further to simulate the human articulatory system, the vocal cord orifice is represented by a series circuit connection of a variable resistor and a variable inductor. A vocal tract excitation signal is produced by supplying a signal representative of subglottal pressure to the circuit and varying the values of the circuit elements in dependence upon glottal area. The glottal area information is obtained for this system from photographs of the vocal cords of a human subject as he utters a sound. Obviously, such a system is limited in application. Further, the system still does not yield a versatile excitation signal and, moreover, requires vocal cord area data as an additional input.

Therefore, although such a system may be advantage-

2

ously employed to analyze certain sounds, it is unsatisfactory for practical use in synthesizing total human speech. Furthermore, difficulty is experienced in developing a vocal tract excitation signal in accordance with a given phonemic sequence.

SUMMARY OF THE INVENTION

Therefore, it is an object of this invention to generate natural-sounding speech sounds.

Another object of the invention is to generate a natural vocal tract excitation signal in response to coded input signals.

Still another object of this invention is to generate, automatically, a vocal tract excitation signal in accordance with physiological parameters of the human glottal system.

In accordance with this invention, these and other objects are accomplished by turning to account certain properties of the human articulatory system. Important among these, is the realization that phonation takes place in the human articulatory system through the vibratory action of the vocal cords. Several forces, acting upon the vocal cords, cooperate to vary the vocal cord orifice, or glottal area, thereby controlling air flow into the vocal tract. The air flow through a properly configured tract gives rise to audible speech. In like manner, control signals developed to represent the momentary glottal area and the driving air force for a particular sound, may be used to control the synthesis of artificial speech.

I have discovered that a function proportional to the glottal area may be generated automatically by representing the vocal cords as a second order oscillatory system that is driven by a nonlinear forcing function. For example, symmetrical vocal cords may be represented as a mechanical oscillator of mass M , spring constant K , and viscous damping B , which is driven by a forcing function F . Mass M is fixed at a value representative of human cords, spring constant K is selected to correspond to vocal cord tension, and viscous damping B is valued to represent rubbing and collision of the cords. Forcing function F is chosen to represent orifice inlet and outlet pressures acting upon the vocal cord surface area. These pressures are related to the subglottal pressure and the resultant Bernoulli pressure with the orifice.

By supplying signals representative of subglottal pressure to an oscillatory system of this sort in accordance with parameter values dictated by sequences of phonemes, word groups, or the like, a glottal area function is developed which may be used, in turn, to generate a vocal tract excitation signal capable of controlling the synthesis of natural-sounding speech.

Accordingly, in the practice of this invention, a vocal tract excitation signal is generated by representing the human vocal cord orifice by a series circuit connection of two variable resistors and a variable inductor, and by controlling the values of the circuit components in accordance with variations in glottal area.

Therefore, a feature of the present invention is the determination of pitch inflection, glottal waveform, source-system interaction, stress, and irregular temporal changes in signal detail according to the physiological parameters of subglottal pressure, vocal cord tension and vocal tract shape.

These and other objects and advantages of the invention will be more fully understood from the following detailed description of an illustrative embodiment thereof taken in connection with the appended drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG 1 illustrates a speech synthesizer that employs the present invention;

FIG. 2 shows in block schematic form the details of the excitation signal generator of FIG. 1;

FIG. 3A shows diagrammatically a mechanical model of the vocal cords for elastic collision;

FIG. 3B shows a mechanical model of the vocal cords for a viscous collision;

FIG. 3C shows diagrammatically another view of the model of FIG. 3B;

FIG. 4 shows graphically the distribution of pressure within the vocal cord orifice;

FIG. 5A shows graphically a typical subglottal pressure characteristic;

FIG. 5B shows graphically a typical fundamental frequency characteristic;

FIG. 6A shows graphically variations in glottal area and an excitation signal during an elastic collision of the vocal cords;

FIG. 6B shows graphically variations in glottal area and an excitation signal during a viscous collision of the vocal cords;

FIG. 7 shows in block schematic form the details of the glottal area function generator of FIG. 2; and

FIG. 8 shows in block schematic form the details of a variable resistor 21 suitable for use in the apparatus of FIG. 2.

DETAILED DESCRIPTION OF THE INVENTION

A spoken word is composed of a sequence of linguistic elements called "phonemes." The inflection given to a word, or a group, often determines meaning. Certain words are thus usually stressed when spoken, in the fashion indicated by punctuation in writing. For the production of synthetic speech, phonemes, or sequence of phonemes, together with suitable stressing must be specified, preferably in a form directly acceptable by a synthesizer.

FIG. 1 shows in block schematic form apparatus which responds to a stressed phonemic input to develop signals useful for producing synthetic speech. Phoneme symbol sequence generator 10 is utilized to produce appropriate signals for initiating synthesis of a spoken word. In generator 10, symbols corresponding to the phonemes of word groups and their associated stress (punctuation) are stored, for example, on a tape in coded form. Upon request, this information is supplied, preferably, although not necessarily, in digital form, to vocal tract area function generator 11, subglottal pressure function generator 12, and vocal cord tension function generator 13.

Operation of phoneme symbol sequence generator 10 is conventional; the generator may be of any desired form. For example, it may be of the type described in detail in Dudley-Harris Pat. 2,771,509, or in Gerstman-Kelly Pat. 3,158,685. If so desired, either of the codes described in the respective patents may be utilized in the practice of the present invention.

Once a stored word group, that is, a phonemic sequence, has been read out in coded form, it is necessary to generate additional signals that are utilized to produce synthetic speech. In accordance with this invention, each phoneme of a particular sequence is employed to control the generation of a plurality of vocal tract area control signals, Av_1 through Av_n , which represent the changes in area of respective segments of the vocal tract, and a vocal tract excitation signal U_g . These signals are sufficient to energize and control vocal tract synthesis apparatus 30.

Area signals Av_1 through Av_n for controlling the synthesis of voiced sounds by vocal tract synthesizer 30 are generated by vocal tract area function generator 11. In order to obtain natural sounding speech, it is preferable that apparatus 11 generate vocal tract area control signals in accordance with certain physiological characteristics of the human vocal tract. Physiological information is therefore used to obtain the proper interpolation between the area control signals generated for one phoneme in a particular sequence and the area control signals generated for the next phoneme in the sequence. This in-

formation is included in the area control signals Av delivered by generator 11 to synthesizer 30.

Function generator 11 also develops signals which represent the source (lv_1) and terminating (lv_2) lengths of the vocal tract. These data are used to control corresponding adjustments in synthesizer 30. Finally, a signal which denotes the duration t of the entire sequence of area control signals for a particular word group is also generated by apparatus 11. Use of the time duration signal will be described later in connection with descriptions of sub-glottal pressure generator 12 and vocal cord tension generator 13.

Vocal tract area function generating apparatus which provides all of the necessary synthesizer control signals, and which is satisfactory for use in the practice of the invention, is described in detail in a copending application of C. H. Coker, filed Aug. 29, 1967, Ser. No. 664,129.

In developing natural sounding synthetic speech, it is additionally necessary to generate a suitable vocal tract excitation signal. Thus, in the practice of this invention, an excitation signal is generated automatically that is related to a corresponding word group. Subglottal pressure function generator 12 and vocal cord tension function generator 13 are used for this purpose.

In speaking a particular word group, or phrase, one gives it a certain inflection. For example, it may be either a declarative, a question of an exclamation, which one intends another to perceive. Therefore, in order to produce natural sounding speech through synthesis, it is necessary to generate signals which exhibit these characteristics. Thus, information is required in addition to the segmental signals of the phoneme sequence. Consequently, subglottal pressure function generator 12 and vocal cord tension function generator 13 are utilized to supply information which is later used in the generation of the vocal tract excitation signal in excitation generator 20, to be explained later in connection with the details as shown in FIG. 2. Typical examples of this supersegmental information are shown in FIGS. 5A and 5B. Subglottal pressure and fundamental frequency (vocal cord tension is proportional to the square of fundamental frequency), corresponding to a particular word group or phrase, are shown.

This supersegmental information may be stored in any of a number of ways for future use, as required. For example, the magnitude of the respective information relating to each of n segments of time Δt , may be established, for example, by the adjustment of a potentiometer. Thus, the subglottal pressure signal P_s and vocal cord tension signal $T(t)$ can be generated by sweeping through n potentiometers during time duration t of the associated phonemic sequence making up the desired word group. A signal identifying time duration t is supplied from the vocal tract area function generator 11, to both generators 12 and 13, and the associated phonemic sequence, in coded form, is supplied to both generators from phoneme symbol sequence generator 10. This information, i.e., subglottal pressure and vocal cord tension, may also be stored in digital form, for example, in an associative storage system which permits certain information, corresponding to the particular input phonemic sequence plus punctuation, to be read out. The digital output from such a system may readily be converted to analog form in any of a number of ways well known in the digital-to-analog conversion art.

Apparatus for accumulating the necessary physiological data, such as subglottal pressure and fundamental frequency relating to any desired word group or phrase, is disclosed in chapter 4 of *Intonation, Perception and Language*, by Philip Lieberman, MIT Press 1967. Diagrams representative of subglottal pressure and fundamental frequency related to certain word groups are also to be found in the Lieberman reference. These data are stored for future use by function generators 12 and 13.

The outputs of generators 12 and 13, i.e., P_s and $T(t)$, respectively, are supplied to excitation signal generator 20

5

where, in accordance with this invention, a vocal tract excitation signal, U_g , is generated. Details of generator 20 and its operation are discussed below with reference to FIG. 2.

Excitation signal U_g , vocal tract area signals Av_1 through Av_n , and vocal tract length control signals lv_1 and lv_2 , are applied to vocal tract synthesizer 30. Each of the applied vocal tract signals controls the value of a circuit element in the synthesizer which represents an element of the human vocal tract. Excitation signal U_g represents the air flow into the human vocal tract. Synthesizer 30, of known construction, responds to the applied control signals and develops artificial speech sounds which may be used to energize a transducer such as loudspeaker 32.

In addition to voiced output signals generated in response to vocal tract and excitation control signals, an unvoiced signal output must be generated. This is accomplished by inserting an appropriate noise source within vocal tract synthesizer 30 in a fashion also well known in the art. A suitable transmission line type of synthesizer which may be used in the practice of the present invention is discussed in detail in "Dynamic Analog Speech Synthesizer," by George Rosen in The Journal of the Acoustical Society of America, March 1958.

FIG. 2 shows in block schematic form the detail of excitation generator 20. Apparatus is utilized including: an orifice circuit which comprises variable resistor 21, variable resistor 22, and variable inductor 23; delay circuit 25; and glottal area function generator 24.

The vocal cord orifice, or glottal impedance, may be approximated by a series combination of resistors and an inductor that vary in value in proportion to the glottal area and excitation signal. Accordingly:

$$R_{21}(t) = k_2 |U_g| A_g^{-2} \quad (1)$$

$$R_{22}(t) = k_3 A_g^{-3} \quad (2)$$

$$L_{23}(t) = k_4 A_g^{-1} \quad (3)$$

where A_g represents glottal area, and k_2 , k_3 , and k_4 are constants. A more complete discussion of the theoretical considerations and derivations of above Equations 1, 2, and 3 related to the vocal cord orifice can be found in section 3.52 of Speech Analysis, Synthesis, and Perception, previously cited.

Variable resistor 21, discussed later in conjunction with FIG. 8, includes apparatus for controlling its value in response to a signal related to Equation 1. Variable resistor 22 (FIG. 2) includes appropriate multiplying and scaling circuitry to obtain a signal related to Equation 2. The variable resistor can be either a rheostat driven by a servomechanism or an appropriately biased field effect transistor. Variable inductor 23, again, may be one which is controlled by a servomechanism and incorporates the proper amplifying means for scaling the glottal area signal to meet the conditions of Equation 3. Glottal area function generator 24, described in detail with reference to FIG. 7, develops automatically, in accordance with the invention, area function, A_g , which, in turn, is applied to the orifice circuit components.

To generate glottal area function signal A_g , certain physiological characteristics of the human glottal system must be considered. I have discovered that the vocal cords may be represented as a second order mechanical system that is driven by a nonlinear forcing function. A typical example of such a mechanical system is shown in FIG. 3A. In this instance, it is assumed that the vocal cords are symmetrical and that there is an elastic collision between them. The elastic collision is simulated by mass M_1 striking block 70. Mass M_1 is set at a value for a typical set of human vocal cords, B_1 is a viscous damping factor selected to simulate cord rubbing, and K_1 is the equivalent of vocal cord tension. The mass is shown, in FIG. 3A, at the phonation neutral position, that is, the position of the vocal cords just before commencing phonation. In this position, displacement is equal to $x=0$. An oscillating

6

system of this type may be expressed mathematically as follows:

$$M_1 \frac{dx^2}{dt^2} + B_1 \frac{dx}{dt} + K_1 x = F(t) \quad (4)$$

and

$$F(t) = \frac{1}{2} (P_1 + P_2) (w \cdot d) \quad (5)$$

where $F(t)$ represents a forcing function, $(w \cdot d)$ is the vocal cord surface area, and P_1 and P_2 are the inlet and outlet pressures, respectively, of the vocal cord orifice. Pressures P_1 and P_2 are related to the subglottal pressure P_s and Bernoulli pressure P_B in the vocal cord orifice according to the graph, as shown in FIG. 4, which is a diagram of pressure versus distance through the vocal cord orifice. Thus, as shown in FIG. 4, at distance 0 the vocal cord orifice pressure P_1 is $(P_s - 1.37 P_B)$ and, at distance d the vocal cord orifice, output pressure P_2 is equal to $(-0.5 P_B)$. The pressures can be negative in magnitude due to the Bernoulli pressure in the orifice. Thus, substituting appropriate values from the graph of FIG. 4 for P_1 and P_2 in Equation 5

$$F(t) = k_5 (P_s - k_6 P_B) \quad (6)$$

where

$$k_5 = (w \cdot d) / 2, \quad k_6 \approx 1.87$$

and

$$P_B = k_7 |U_g|^2 A_g^{-2} \quad (7)$$

Therefore, glottal area may be obtained by solving Equation 4 for displacement x and substituting in

$$A_g = A_{g0} + (w \cdot x) \quad (8)$$

where A_{g0} represents glottal area with the vocal cords in the neutral position, w represents the width of the vocal cord orifice, and $A_g = U_g = 0$ for $x = x_c$, where

$$x_c = -A_{g0} / w$$

Applying a signal representative of subglottal pressure P_s , as shown in FIG. 5A, and vocal cord tension $T(t)$, which is proportional to the square of fundamental frequency, as shown in FIG. 5B, to glottal excitation source 20 (FIG. 2), the glottal excitation signal is generated as follows: an initial value neutral glottal area A_{g0} is supplied to control variable resistors 21 and 22 and variable inductor 23 in order to produce an initial current value for excitation signal U_g . Area signal A_{g0} is delayed by time delay circuit 25, which may be simply a conventional delay line, and is fed back as delayed signal A_{g0}' to glottal area function generator 24. The delay is necessary so that the proper value of $F(t)$ is generated. Thus, feeding back A_{g0} and the corresponding value of U_g , $F(t)$ is generated, from which displacement x is obtained, and subsequently, a new value of A_g . This process is iterated, thereby obtaining a continuous vocal tract excitation signal. Generator 20 is self-oscillating, requiring only the input of signals representative of subglottal pressure and vocal cord tension plus the loading effect of vocal tract synthesizer 30 to produce automatically an excitation signal U_g . Such a signal is shown in FIG. 6A where glottal area A_g and excitation signal U_g are shown with respect to time.

A more accurate mechanical representation of the human vocal cords is shown in FIG. 3B and FIG. 3C. This approximation again assumes symmetrical vocal cords but in this instance the collision of the cords is assumed to be viscous, i.e., inelastic, or semielastic. The collision of mass M_1 with dashpot 71 creates an additional damping B_2 which causes the glottal area A_g and excitation signal U_g to remain at a zero value for a portion of each cycle. The resultant variations in the glottal area and excitation signals, as shown in FIG. 6B, are believed to be a more accurate representation of the corresponding human func-

tions. The mathematical representation of such a system for a viscous collision is as follows:

$$M_1 \frac{dx^2}{dt^2} + B_1 \frac{dx}{dt} + K_1 x = F(t) + F_o \quad (9)$$

where

$$F_o = \begin{cases} 0 & \text{for } x \geq x_o \\ -B_2 \frac{dx}{dt} & \text{for } x < x_o \end{cases}$$

$$F(t) = k_5(P_s - k_6 P_B) \quad [6]$$

$$P_B = k_7 |U_g|^2 A_g^{-2} \quad [7]$$

and $A_g = U_g = 0$, for $x \leq x_o$, where $x_o = A_{go}/w$. The sequence for generating the glottal area signal and, in turn, the excitation signal, is as discussed above in relation to Equation 4.

FIG. 7 shows in block schematic form the details of glottal area function generator 24. The delayed area function signal A_g' is squared in squaring circuit 40, scaled in amplifier 41, and applied to one input of divider 42. Excitation signal U_g , corresponding to A_g' , is squared in squarer 43, scaled in amplifier 44, and applied to the other input of divider 42. The resultant signal $k_6 P_B$ is applied to one input of subtracter 45. Subglottal pressure signal P_s is applied to the other input of subtracter 45 where the applied signals are accordingly subtracted and scaled to generate a signal which represents forcing function $F(t)$. The forcing function signal and a signal representing vocal cord tension $T(t)$ are fed into computing network 50, preferably an analog computer, where they are utilized to generate a displacement signal x . Computing circuits of the type utilized in the present invention are discussed in chapter 5 of Analog Computation, by Albert S. Jackson, McGraw-Hill 1960. Particularly, the instances of elastic and viscous collisions are discussed at pages 200 through 203 in the Analog Computation reference. In computing circuit 50, the vocal cord tension signal $T(t)$ is used to control the coefficient K_1 in Equations 4 and 9. This variation is in accordance with the voiced fundamental frequency f_o shown in FIG. 5B as follows:

$$T \approx K_1$$

$$F_o \text{ (mech.)} = \frac{1}{2\pi} \sqrt{K_1/M_1}$$

Thus,

$$T = K_1 = 2\pi F_o^2 / M_1$$

As a first approximation, the voiced fundamental frequency f_o (voiced) is

$$f_o \text{ (voiced)} \approx k_8 F_o \text{ (mech.)}$$

therefore,

$$T(t) = K_1(t) = k_9 (f_o(t))^2 \quad (10)$$

The coefficient relating to vocal cord tension, that is, K_1 , can be varied in accordance with principles well known in the analog computation art. The resultant output from network 50 is displacement signal x . Displacement signal x is scaled by the factor w in amplifier 51 and added to area function signal A_{go} in summing amplifier 52. The resultant sum represents the desired glottal area A_g .

Level detector 53 and switching transistors 54 and 55 are utilized to ensure that the glottal area signal and the excitation signal are equal to 0 for values of $x < x_o$. This feature may be employed, if desired, in addition to the limiters usually incorporated in computing network 50.

FIG. 8 shows in block schematic form the details of variable resistor 21 (of FIG. 2). According to the conditions of Equation 1 glottal area signal A_g is squared in squaring network 60 and the resultant is applied to one input of divider 61. Excitation signal U_g is applied to circuit 62, such as a full wave rectifier, where its absolute magnitude $|U_g|$ is established. This value is applied to the second input of divider 61. The output of the divider 61,

$k_2 |U_g| A_g^{-2}$ controls variable resistor 63. Variable resistor 63 may be any of a number of types, for example, it can be a rheostat that is driven by a servomechanism, or a field effect transistor that is appropriately biased to be varied by an input signal corresponding to Equation 1.

The above-described arrangements are, of course, merely illustrative of the application of the principles of this invention. Numerous other arrangements may be devised by those skilled in the art without departing from the spirit and scope of the invention. For example, the vocal cords may be represented as a dual mechanical oscillatory system in which each cord is simulated by an individual mechanical oscillator or further, each cord may be simulated by a matrix of interconnected masses and springs which are driven by an appropriate forcing function.

What is claimed is:

1. Apparatus for generating a vocal tract excitation signal which comprises,
 - means for generating signals representative of an ordered phonemic sequence,
 - means responsive to said signals representative of said phonemic sequence for generating a control signal representative of subglottal pressure,
 - means responsive to said signals representative of said phonemic sequence for generating a control signal representative of vocal cord tension, and
 - generating means responsive to said subglottal pressure and vocal cord tension control signals for generating a vocal tract excitation signal.
2. Apparatus for generating a vocal tract excitation signal which comprises, in combination,
 - means for generating coded signals representative of phonemes in a desired sequence,
 - first means responsive to said coded signals for deriving a signal representative of subglottal pressure for the duration of said sequence,
 - second means responsive to said coded signals for deriving a signal representative of vocal cord tension for the duration of said sequence,
 - function generator means responsive to said pressure and tension signals for generating a signal representative of glottal area, and
 - controllable impedance means responsive to said area signal for varying a signal applied to a vocal tract synthesizer.
3. Apparatus as defined in claim 2 further including means for delaying said glottal area signal a predetermined interval and wherein said function generator means comprises,
 - means responsive to said delayed glottal area signal and said signal applied to said vocal tract synthesizer for generating a signal representative of the Bernoulli pressure within the vocal cord orifice,
 - means responsive to said signal representative of subglottal pressure and to said Bernoulli pressure signal for generating a signal representation of a forcing function $F(t)$,
 - computing means supplied with said signal representation of forcing function $F(t)$ and with a signal representative of said vocal cord tension for generating a glottal displacement signal x , said computing means being preset in accordance with preestablished vocal cord mass and viscous damping parameters, and
 - means responsive to said glottal displacement signal x for generating a signal representative of glottal area.
4. Apparatus for generating an excitation signal for energizing a vocal tract synthesizer which comprises, in combination:
 - means for generating a signal representative of a sequence of phonemic symbols,
 - first means responsive to said phonemic sequence sig-

nal for generating a signal representative of subglottal pressure,
 second means responsive to said phonemic sequence signal for generating a signal representative of vocal cord tension,
 means responsive to said pressure and tension signals for generating a controlling signal, and
 means responsive to said controlling signal for varying signals supplied as excitation to a vocal tract synthesizer.
 5
 5. Apparatus as defined in claim 4 wherein said phonemic sequence signal generator comprises:
 means for reading said phonemic sequences from a tape, and
 means for translating said sequence in coded signal form.
 10
 6. Apparatus as defined in claim 4 wherein said first signal generating means comprises:
 means for storing information representative of subglottal pressure, and
 means for selectively reading out said information in response to said phonemic sequence signal to produce a signal.
 20
 7. Apparatus as defined in claim 4 wherein said second generating means comprises:
 means for storing information representative of vocal cord tension, and
 means for selectively reading out said information in response to said phonemic sequence signal to produce a signal.
 25
 8. Apparatus as defined in claim 4 wherein said control signal means comprises,
 means responsive to a signal representative of said glottal area and to said excitation signal for generating a signal proportional to the Bernoulli pressure in the vocal cord orifice,
 means responsive to said subglottal pressure signal and to said Bernoulli pressure signal for generating a forcing function signal $F(t)$,
 computing means for generating a glottal displacement signal x in response to said forcing function and said vocal cord tension, and
 means for generating a glottal area signal A_g in response to said glottal displacement signal.
 30
 9. Apparatus as defined in claim 4 wherein said varying means comprises a series of circuit including: first

controllable resistive means, second controllable resistive means, and controllable inductor means.

10. In combination:

means for generating coded signals representative of an ordered sequence of phonemes,
 storage means responsive to said coded signals for generating, for each sequence of phonemes, signals representative of subglottal pressure and vocal cord tension,
 means responsive to said subglottal pressure and vocal cord tension signals for generating signals representative of subglottal area corresponding to said sequence phonemes,
 a vocal tract synthesizer, and
 controllable impedance means supplied with said subglottal pressure signal and responsive to said glottal area signal for controlling the excitation potential applied to said vocal tract synthesizer, said controllable impedance means including a first variable resistor, a second variable resistor and a variable inductor,
 said subglottal area signal generating means including means for delaying said glottal area signal a predetermined interval,
 means responsive to said delayed glottal area signal and to said excitation potential for generating a signal proportional to the Bernoulli pressure within the vocal cord orifice,
 means responsive to said subglottal pressure signal and said Bernoulli pressure signal for generating a signal representative of a forcing function $F(t)$.
 computational means responsive to said forcing function signal $F(t)$ and said vocal cord tension signal $T(t)$ for generating a glottal displacement signal x , said computational means being preset in accordance with preestablished vocal cord mass and viscous damping parameters,
 means responsive to said glottal displacement signal x for generating said glottal area signal A_g , and
 means for limiting said glottal area signal to predetermined values.

No references cited.

45 KATHLEEN H. CLAFFY, Primary Examiner
 R. P. MYERS, Assistant Examiner