

12

EUROPEAN PATENT SPECIFICATION

45 Date of publication of patent specification: **08.06.88**

51 Int. Cl.⁴: **G 10 L 5/04**

21 Application number: **84305918.9**

22 Date of filing: **30.08.84**

54 **Speech synthesis apparatus.**

30 Priority: **31.08.83 JP 157719/83**

43 Date of publication of application:
02.05.85 Bulletin 85/18

45 Publication of the grant of the patent:
08.06.88 Bulletin 88/23

84 Designated Contracting States:
DE FR GB

58 References cited:
GB-A-1 592 473

EDN - ELECTRICAL DESIGN NEWS, vol. 25, no. 14, August 1980, pages 99-103, Denver, US; E. TEJA: "Versatile voice output demands sophisticated software"

ICASSP 82 - PROCEEDINGS OF THE IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 3rd-5th May 1982, Paris, FR, vol. 3, pages 1589-1592, IEEE, New York, US; D.H. KLATT: "The klattalk text-to-speech conversion system"

73 Proprietor: **KABUSHIKI KAISHA TOSHIBA**
72, Horikawa-cho Saiwai-ku
Kawasaki-shi Kanagawa-ken 210 (JP)

72 Inventor: **Watanabe, Sadakazu**
708-132, Ohzenji Asao-ku
Kawasaki-shi Kanagawa-ken (JP)
Inventor: **Nomura, Norimasa**
6 Tkaenomaru Naka-ku
Yokohama-shi Kanagawa-ken (JP)

74 Representative: **Kirk, Geoffrey Thomas et al**
BATCHELLOR, KIRK & EYLES 2 Pear Tree Court
Farringdon Road
London EC1R 0DS (GB)

59 References cited:
ICASSP 80 - PROCEEDINGS OF THE IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 9th-11th April 1980, vol. 2, pages 576-579, IEEE, New York, US; J. BERNSTEIN et al.: "Unlimited text-to-speech system: description and evaluation of a microprocessor based device"

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European patent convention).

Description

One of the known methods for transforming character strings into synthetic speech is "synthesis by rule". In this method, a character string is first transformed into a sequence of phonemes. Next, the prosodic parameters (the duration, pitch and power) of each phoneme are determined, and speed segments having those parameters are selected from a library of the spectral envelopes of such speech segments. Finally, the phoneme sequence and the parameters are provided to a well-known speech synthesizer which assembles the segments into synthetic speech, adjusting the parameters and connecting them into more-or-less natural speech. Although it is possible to synthesize speech using standard values of each parameter for each phoneme, it is also possible to vary the duration of those phonemes which consist of a consonant-vowel (CV) combination. When one of these variable phonemes is encountered in a phoneme string, its duration may be modified (changed from the standard value) by considering the influence of the phoneme immediately before or after the variable phoneme in the phoneme string. However, even when the duration of CV phonemes is modified in accordance with the influence of influencing phonemes immediately adjacent the variable phoneme, the synthetic speech produced is rather unnatural and unclear.

A modification rule is described in EDN, Vol. 25, No. 14, August 1980 pp 99—103, Teja.

One object of this invention is to produce synthetic speech of such clarity and high quality that it is very nearly natural speech.

Another object of the invention is to produce synthetic speech in which the prosodic parameters of variable phonemes are modified as a function of the existence of influencing phonemes at locations in the phoneme string other than the two locations immediately adjacent the variable phoneme (as a function of a non-influencing phoneme in a non-adjacent mora).

The invention is based on the careful observation of natural speech by the present inventors. They found, for example, that the duration of a variable phoneme consisting of a double consonant is influenced more by the kind of phoneme which exists in the phoneme string in a location one phoneme removed from the double consonant than by the kind of phoneme which exists immediately adjacent the double consonant in the phoneme string. They observed that, in Japanese speech, a double consonant lasts several milliseconds longer when a "prolonged sound" or a syllabic nasal "N" is at a location in the phoneme string two morae subsequent to the double consonant.

The invention comprises five main elements. A character string translator accepts encoded character strings from a device, such as a keyboard, which is capable of inputting character strings electronically. Using a word memory, the character string translator converts each word in the input character string into a phoneme string corresponding to the characters in the word. Next, a variable phoneme detector detects those phonemes in the string the values of whose prosodic parameters may be modified due to the existence of an influencing phoneme at a location in the phoneme string which is at least one mora removed from the location of the variable phoneme. If variable phonemes are detected in the phoneme string, a search is made by an influencing phoneme detector for influencing phonemes at the location in the phoneme string indicated by the variable phoneme detector. The variable phoneme detector is associated with a variable phoneme memory which stores, along with each variable phoneme, the predetermined location at which an influencing phoneme will influence the value of each parameter of the variable phoneme. If an influencing phoneme is detected at the appropriate location in relation to the variable phoneme the influencing phoneme detector will output data representative of a modification in the value of a selected parameter (duration, pitch or power) of the variable phoneme. The phoneme string is then delivered to a parameter value determining unit, which stores standard values of the parameters for all phonemes. Standard values may be modified in response to the modification data supplied by the influencing phoneme detector. Finally, the phoneme string, parameter values, and modification data are supplied to a well-known parametric synthesizer, which assembles them into synthetic speech. Of course, since this invention is an electronic apparatus, it does not perform operations on "characters" or "phonemes" but rather on electrical codes representing characters and phonemes. This fact will be implicitly recognized throughout the specification.

According to one aspect of the invention there is provided an apparatus for creating synthetic speech from a character string, comprising: translating means for translating the character string into a phoneme string, parameter value determining means for storing a value of a selected parameter of each phoneme in the phoneme string, and parametric synthesizer means for assembling the phoneme string into synthetic speech using the stored values, characterised by variable phoneme detecting means coupled to said translating means for detecting a variable phoneme comprising: a double consonant (Q) a syllabic nasal (N) or a prolonged sound (L) in the phoneme string, the value of the selected parameter being a function of the identity of an influencing phoneme in the phoneme string; and influencing phoneme detecting means coupled to said variable phoneme detecting means for determining the identity of the influencing phoneme which is also a double consonant (Q), a syllabic nasal (N) or a prolonged sound (L), said parameter value determining means being responsive to the result of the determination for modifying the stored value of the selected parameter, said influencing phoneme being located at a predetermined location, non-adjacent to the variable phoneme in the phoneme string.

According to another aspect of the invention there is provided a method of creating synthetic speech

from a character string, comprising: translating the character string into a phoneme string, storing a value of a selected parameter of each phoneme in the phoneme string, and assembling the phoneme string into synthetic speech using the stored values, characterised by detecting a variable phoneme comprising a double consonant (Q), a syllabic nasal (N) or a prolonged sound (L) in the phoneme string, the value of the selected parameter being a function of the identity of an influencing phoneme in the phoneme string; and determining the identity of the influencing phoneme which is also a double consonant (Q), a syllabic nasal (N) or a prolonged sound (L) and responsive to the result of the determination modifying the stored value of the selected parameter, said influencing phoneme being located at a predetermined location non-adjacent to the variable phoneme in the phoneme string.

Brief description of the drawings

Figure 1 is a block diagram of one embodiment of the improved speech synthesis apparatus.

Figure 2 is a schematic diagram showing in more detail the character string translator of Figure 1.

Figure 3 is a schematic diagram showing in more detail the variable phoneme detector of Figure 1.

Figure 4 is a schematic diagram showing in more detail the influencing phoneme detector of Figure 1.

Figure 5 is a schematic diagram showing in more detail the parameter value determining unit of Figure 1.

Figure 6 is a block diagram of another embodiment of the improved speech synthesis apparatus.

Detailed description of the preferred embodiment

Fig. 1 is a block diagram of the improved speech synthesis apparatus. An encoded character string is input from a device (not shown) such as a keyboard having character keys, a memory device which stores character strings (such as used in a word processor), or a communication device receiving character strings through communication lines.

Character string translator 1 translates the character codes making up the character string into phoneme codes representing a phoneme string, using word memory 2. The phoneme string is supplied to variable phoneme detector 3, which detects variable phonemes. In a Japanese phoneme string, these variable phonemes include the double consonant (hereinafter indicated by "Q"), syllabic nasal "N", and prolonged sound (hereinafter indicated by "L").

Such variable phonemes are stored in variable phoneme memory 4 which is used by detector 3 in the detection of variable phonemes. As is described later, if a variable phoneme is detected, it is associated with attribution data from which parameters value modifications may be determined. The output of detector 3 includes information regarding the predetermined location at which an influencing phoneme must be found in order to influence the value of a parameter (the duration, in the preferred embodiment) of the detected variable phoneme.

Influencing phoneme detector 5 determines whether there exists an influencing phoneme, such as a double consonant, syllabic nasal "N" or prolonged sound at the location indicated by detector 3.

Parameter value determining unit 7 stores the normal values of parameters such as the duration of the phoneme, its pitch or its power. If both the variable phoneme and a corresponding influencing phoneme are detected unit 7 outputs a modified value of the given parameter to parametric synthesizer 8, which may be any well-known synthesizer, for example formant-type, Parcore-type, and Ceptstrum-type synthesizer.

Fig. 2 shows in more detail the character string translator 1 of Fig. 1. A character string (which normally constitutes one word) enters input register 101 under control of input control circuit 102. Read control circuit 103 supplies an initial address to word memory 2, and a word is read out into register 104. Word memory 2 stores a plurality of words (in segment 104A) together with the corresponding phoneme strings into which the words are translated (in segment 104B). The word from segment 104A is supplied to comparator 100 and compared with the content of input register 101. If the input character string in input register 101 is not the same as the character string from segment 104A, comparator 105 produces a signal on line 106. Responding to the signal on line 106, read control circuit 103 increments its internal counter (not shown) and provides the incremented address to word memory 2, so that the next word in memory 2 (together with the corresponding phoneme string) is read out into register 104. These operations are repeated until comparator 105 detects identity between the input character string and the character string retrieved from memory 2.

When identity is detected between the input character string in input register 101 and the character string in segment 104A of register 104, the comparator produces a signal on line 107, causing the phoneme string in segment 104B to be written into phoneme string memory 108. At the same time, input control circuit 102 causes the next character string to enter input register 101, and read control circuit 103 supplies the initial address to word memory 2 in preparation for the next translation operation.

Fig. 3 shows in more detail the variable phoneme detector 3 of Fig. 1. Input control circuit 301 accesses phoneme string memory 108 and successively transfers phonemes into buffer 302. Most Japanese speech consists of four morae, as follows (a "mora" is a basic unit of time in speech; it may contain one or more phonemes):

- (a) consonant-vowel combination ("CV")
- (b) an independent vowel, or a prolonged sound (symbolized as "L")
- (c) a double consonant (symbolized as "Q")

(d) a syllabic nasal "N"

Read control circuit 303 supplies an initial address to variable phoneme memory 4, which stores variable phonemes in association with the predetermined location in the phoneme string at which an influencing phoneme will influence the value of the duration of each variable phoneme. Variable phoneme data from memory 4 is written into register 304, which consists of a phoneme segment 304A and a relative location segment 304B. The relative location data in relative location segment 304B consists of at least one location in the phoneme string, relative to the variable phoneme, where an influencing phoneme may be located to influence the duration of the variable phoneme. For example, if the variable phoneme is a double consonant, "Q", its relative location data may be "+2", which means that the duration of the double consonant is influenced by an influencing phoneme located two morae later in the phoneme string. Comparator 305 compares the phonemes in phoneme buffer 302 and the phoneme segment 304A of register 304. If they are not identical, comparator 305 produces a signal on line 306. Responding to the signal on line 306, read control circuit 303 increments its internal counter (not shown) and provides the incremented address to variable phoneme memory 4, so that the next variable phoneme in memory 4 is read out into register 304. These operations are repeated until comparator 305 detects identity between the phoneme in buffer 302 and the phoneme retrieved from memory 4 as stored in segment 304A, or until all variable phonemes in memory 4 have been retrieved. When identity is detected between the phoneme in buffer 302 and the phoneme in segment 304A, comparator 305 produces a signal on line 307, causing the phoneme from buffer 302 and the relative location data in segment 304B (through selector 308) to be written into memory 309. The selector 308 normally selects the relative location data from segment 304B; however, if identity is never detected by comparator 305, read control circuit 303 outputs an END signal on line 310, causing selector 308 to select "O" as the relative location data, indicating that the phonemes are not variable phonemes.

Fig. 4 shows in more detail the influencing phoneme detector 5 of Fig. 1. Read control circuit 501 controls counter 502, whose content indicates an address memory 309. The present address in counter 502 is supplied to memory 309 through adder 503 which normally adds "O" to the present address. The data read out from memory 309 are provided to register 505 through selector 504 which normally selects register 506. The data consists of phonemes (in segment 505A) and relative locations (in segment 505B). The relative location information in segment 505B is supplied to zero detector which produces a signal on line 507 when the relative location information in segment 505B is "O" or on line 508 when the relative location information is not "O". In the latter case, the signal on the line 508 is provided to read control circuit 501 to prevent the incrementing of counter 502, to the selector 504 to select register 506, and to read control circuit 510 to supply an initial address to parameter value variation memory 6. That is, when the relative location information is not "O", the relative location is added to the present address in counter 502 by adder 503, and the sum is supplied as an address to memory 309 so that the corresponding data is read from memory 309 and stored in register 506 via selector 504.

Parameter value variation memory 6 is accessed by read control circuit 510 and supplies to register 512 data consisting of a phoneme, in segment 512A (an influencing phoneme) and, in segment 512B, attribution data from which the change in duration of the variable phoneme may be determined. Comparator 513 compares the phoneme in segment 512A with that in segment 506A, producing a signal on line 514 if it detects identity and on line 515 otherwise. Responding to the signal on line 514, the phoneme in segment 505A and the attribution data in segment 512B are written into a checked data memory 520. On the other hand, the signal on line 515 is provided to read control 510 to increment an inner counter (not shown) and output the next address to memory 511. The comparator compares the phoneme in register 506 with successively retrieved phonemes in register 512. If none of the phonemes in memory 511 is the same as with the phoneme in register 506, read control circuit 510 outputs an AND signal to selector 516. Selector 516 usually selects the attribution data segment of memory 512 as input data but instead selects "O" upon receipt of an END signal. Therefore, the phoneme in segment 505A is written into checked data memory 520 with attribution data equal to "O". When zero detector 506 detects a "O", it outputs a signal on line 507. In this case, the phoneme in segment 505A is not a variable phoneme, so it is written into the checked data memory 520 with attribution data equal to "O".

Fig. 5 shows in more detail the parameter value determining unit 7 of Fig. 1. Data consisting of phonemes in combination with attribution data are successively read from checked data memory 520 into register 701. Parameter value memory 702 stores standard values of the parameters for every phoneme. (It is also possible, instead of storing all the values, to use a parameter calculator according to a phoneme code). A phoneme in segment 701A of register 701 is supplied to parameter value memory 702 as an address; memory 702 then outputs the corresponding parameter value to adder 703. Modifying data memory 704 stores parameter value modifications and outputs them when addressed by attribution data from segment 701B. The standard parameter values from memory 702 and the modification data from memory 704 are added by adder 703 and the sums supplied along with the phoneme string to parametric synthesizer 8 of Fig. 1 for assembly into synthetic speech.

A pitch modifying circuit 710 and a power modifying circuit 720 may be provided for modifying pitch or power data in a similar manner.

TABLE 1

Example	Character string	Phoneme string	Duration of (t_Q) of "Q" phoneme
A.	Kessen	KE/Q/SE/N	$t_Q = t_m + t_a$
B.	Kesseki	KE/Q/SE/KI	$t_Q = t_m$
C.	Ikkan	I/Q/KA/N	$t_Q = t_m + t_a$
D.	Ikkatsu	I/Q/KA/TSU	$t_Q = t_m$

Table 1 shows how the duration of a phoneme is modified. In example A, when the character string "kessen" (a Japanese word meaning "decisive battle") is entered, character string translator 1 translates it into the phoneme string KE/Q/SE/N, where "/" indicates divisions between morae. Variable phoneme detector 3 detects a variable phoneme "Q" in the second location; consequently, influencing phoneme detector 5 should search for an influencing phoneme. In this example, there is an influencing phoneme "N" in the fourth location in the phoneme string, so parameter value determining unit 7 adds a modifying value t_a to the standard double consonant duration t_m , which means that the duration of phoneme "Q" is given by $t_Q = t_m + t_a$.

If character string "kesseki" (meaning "absent") is entered (example B), the translated phoneme string is KE/Q/SE/KI. Variable phoneme detector 3 again detects a variable phoneme "Q" in the second location, but influencing phoneme detector 5 does not detect an influencing phoneme. Therefore, parameter value determining unit 7 determines the duration of the phoneme "Q" as $t_Q = t_m$, the standard double consonant duration.

In Example C, "Ikkan" (meaning "consistently") is translated into the phoneme string I/Q/KA/N including variable phoneme "N" in the fourth location. The duration t_Q of the second phoneme is therefore given by $t_Q = t_m + t_a$. The character string of example D, "Ikkatsu" (meaning "together"), is translated into the phoneme string I/Q/KA/TSU, which includes variable phoneme "Q" in the second location but no influencing phoneme, so that the duration of the second phoneme is $t_Q = t_m$.

The standard duration t_m of a double consonant in Japanese is 170 ms, and additional duration t_a is, for example, 50 ms. According to this embodiment, clear and naturalized speech can be synthesized by considering the influence of a non-adjacent phoneme on the duration of a variable phoneme.

Fig. 6 is a block diagram of another embodiment of this invention. Independent vowel detector 11, neighboring vowel detector 12, and prolonged sound transforming unit 13 are inserted between character string translator 1 and variable phoneme detector 3 of Fig. 1. The remaining elements of Fig. 6 are the same as in Fig. 1, so their descriptions are omitted here.

Most Japanese words consists of four morae; and of these, words having a syllabic nasal "N" in the fourth location appear most frequently. The present inventors have observed that the duration of the syllabic nasal "N" in the fourth location is related to the existence of a syllabic nasal "N" or prolonged sound (L) in the second location.

Independent vowel detector 11, including a comparator (not shown) detects whether the phoneme string includes independent vowels O, U or I. If the phoneme string includes such vowels, neighboring vowel detector 12, also including a comparator, detects the identity of the phoneme immediately preceding the detected vowel. Then prolonged sound transforming unit 13, including a code converter, transforms the detected independent vowel into the prolonged sound L if and only if the combination of the detected independent vowel and the immediately preceding phoneme fall into one of the following categories:

- (a) The detected independent vowel is either U or O and the preceding phoneme is either U or O.
- (2) The detected independent vowel is I and the preceding phoneme is either I or E.

Table 2 shows examples of these cases.

TABLE 2

Example	Original character string	Translated phoneme string	Result of vowel transformation
E.	Kuudoo	KU <u>U</u> DOO	KU <u>L</u> DOL
F.	Ookii	O <u>O</u> KII	OL <u>K</u> IL
G.	Keikei	KEI <u>K</u> EI	KE <u>L</u> KE <u>L</u>

The reason for transformation of vowels into prolonged sounds is that the above independent vowels

become prolonged sounds of the immediately preceding vowels in natural Japanese speech. Moreover, it is helpful, in determining the parameter values of the syllabic nasal or prolonged sound, to consider the influence of a preceding phoneme in the phoneme string.

5

TABLE 3

	Example	Character string	Phoneme string	Duration of fourth phoneme
10	H.	Dangan	DA/N/GA/N	$t_4=t_2$
	I.	Keisan	KE/I/SA/N	$t_4=t_2$
	J.	Tousei	TO/U/SE/I	$t_4=t_2$
15	K.	Nintei	NI/N/TE/I	$t_4=t_2$

Table 3 shows examples of the determination of phoneme duration. In example H, the character string "dangan" (meaning "bullet") includes the syllabic nasal sound "N" in the fourth location. The duration of this sound is influenced by the syllabic nasal sound N in the second location, so the duration of the former is the same as the latter ($t_4=t_2$). These examples indicate that the duration of the syllabic nasal "N" or the prolonged sound L is the same as the duration of the second preceding mora if the phoneme is either at that location the syllabic nasal "N" or the prolonged sound (L).

Although illustrative embodiments of the present invention have been described with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments and that various changes and modifications may be effected therein by one skilled in the art without departing from the scope of the present invention. Specifically, while the preferred embodiment has been illustrative as synthesizing Japanese speech, it is equally applicable to speech of any other language.

30

Claims

1. An apparatus for creating a synthetic speech from a character string, comprising: translating means (1) for translating the character string into a phoneme string, parameter value determining means (7) for storing a value of a selected parameter of each phoneme in the phoneme string, and parametric synthesizer means (8) for assembling the phoneme string into synthetic speech using the stored values, characterised by

variable phoneme detecting means (3) coupled to said translating means for detecting a variable phoneme comprising: a double consonant (Q) a syllabic nasal (N) or a prolonged sound (L) in the phoneme string.

the value of the selected parameter being a function of the identity of an influencing phoneme in the phoneme string; and

influencing phoneme detecting means (5) coupled to said variable phoneme detecting means (3) for determining the identity of the influencing phoneme which is also a double consonant (Q), a syllabic nasal (N) or a prolonged sound (L), said parameter value determining means being responsive to the result of the determination for modifying the stored value of the selected parameter, said influencing phoneme being located at a predetermined location, non-adjacent to the variable phoneme in the phoneme string.

2. An apparatus as claimed in claim 1 wherein the selected parameter is duration.

3. An apparatus as claimed in claim 1 or 2 wherein the translating means (1) has an associated memory (2) for storing a phoneme string corresponding to each of a plurality of words, the translating means having comparing means for outputting the corresponding phoneme string when a stored word matches a portion of the character string.

4. An apparatus as claimed in one of claims 1 to 3 further comprising memory means (4) associated with said variable phoneme detecting means (3) for storing, in association with a variable phoneme, the relative location at which an influencing phoneme will influence the value of the parameter of the stored variable phoneme, the variable phoneme detecting means (3) including comparing means for outputting the associated predetermined location when a stored variable phoneme is detected in the phoneme string.

5. An apparatus as claimed in any preceding claim further comprising memory means (6) associated with the influencing phoneme detecting means (5) for storing, in association with the identity of an influencing phoneme, data representative of a modification in the stored value of the selected parameter, the influencing phoneme detecting means (5) including comparing means for outputting the associated data when the stored influencing phoneme is in the predetermined location.

6. A method of creating synthetic speech from a character string, comprising: translating the character string into a phoneme string, storing a value of a selected parameter of each phoneme in the phoneme string, and assembling the phoneme string into synthetic speech using the

stored values, characterised by

detecting a variable phoneme comprising a double consonant (Q), a syllabic nasal (N) or a prolonged sound (L) in the phoneme string, the value of the selected parameter being a function of the identity of an influencing phoneme in the phoneme string; and

determining the identity of the influencing phoneme which is also a double consonant (Q), a syllabic nasal (N) or a prolonged sound (L) and responsive to the result of the determination modifying the stored value of the selected parameter, said influencing phoneme being located at a predetermined location, non-adjacent to the variable phoneme in the phoneme string.

Patentansprüche

1. Einrichtung zur Schaffung synthetischer Sprache aus einem Eigentümlichkeitsstrang, mit: Übersetzungsmitteln (1) zur Übersetzung des Eigentümlichkeitsstranges in einen Lautstrang, Parameterwert-Bestimmungsmittel (7) zur Speicherung eines Wertes eines ausgewählten Parameters jedes Lautes in dem Lautstrang, und parametrische Synthesiermittel (8) zum Zusammensetzen des Lautstranges in synthetische Sprache unter Verwendung der gespeicherten Werte, gekennzeichnet durch variable Lautfeststellmittel (3), die mit den Übersetzungsmitteln zur Feststellung eines variablen Lautes gekoppelt sind und aufweisen: Einen Doppelkonsonanten-(Q), einen Silbennasal-(N) oder einen langgestreckten Klang (L) in dem Lautstrang,

wobei der Wert des ausgewählten Parameters eine Funktion der Identität eines beeinflussenden Lautes in dem Lautstrang ist und

die beeinflussenden Lautfeststellmittel (5) mit den variablen Lautfeststellmitteln (3) zur Bestimmung der Identität der beeinflussenden Laute gekoppelt sind, die ebenfalls ein Doppelkonsonant-(Q), ein Silbennasal-(N) oder ein langgestreckter Laut (L) sind, wobei die Parameterwert-Bestimmungsmittel auf das Ergebnis der Bestimmung ansprechen und so den gespeicherten Wert des ausgewählten Parameters modifizieren, wobei der beeinflussende Laut an einer vorbestimmten Stelle nicht-benachbart zu dem variablen Laut in dem Lautstrang angeordnet ist.

2. Einrichtung nach Anspruch 1, dadurch gekennzeichnet, daß der ausgewählte Parameter Dauer ist.

3. Einrichtung nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die Übersetzungsmittel (1) einen zugehörigen Speicher (2) zur Speicherung eines Lautstranges entsprechend jedem einer Vielzahl von Worten aufweisen, wobei die Übersetzungsmittel Vergleichsmittel zur Ausgabe des entsprechenden Lautstranges aufweisen, wenn ein gespeichertes Wort mit einem Teil des Eigentümlichkeitsstranges übereinstimmt.

4. Einrichtung nach einem der Ansprüche 1 bis 3, weiter gekennzeichnet durch Speichermittel (4), die den variablen Lautfeststellmitteln (3) zugeordnet sind und in Zuordnung zu einem variablen Laut die relative Lage speichern, bei der ein beeinflussender Laut den Wert des Parameters des gespeicherten variablen Lautes beeinflusst, wobei die variablen Lautfeststellmittel (3) Vergleichsmittel zur Ausgabe des zugehörigen, vorbestimmten Ortes aufweisen, wenn ein gespeicherter, variabler Laut in dem Lautstrang festgestellt wird.

5. Einrichtung nach einem der vorhergehenden Ansprüche, gekennzeichnet durch Speichermittel (6) die den beeinflussenden Lautfeststellmitteln zugeordnet sind und in Zuordnung zu der Identität eines beeinflussenden Lautes Daten speichern, die repräsentativ für eine Änderung des gespeicherten Wertes des ausgewählten Parameters sind, wobei die beeinflussenden Lautfeststellmittel (5) Vergleichsmittel zur Ausgabe der zugehörigen Daten aufweisen, wenn der gespeicherte, beeinflussende Laut an dem vorbestimmten Ort vorhanden ist.

6. Verfahren zur Bildung synthetischer Sprache aus einem Eigentümlichkeitsstrang, bei dem: der Eigentümlichkeitsstrang in einen Lautstrang übersetzt, ein Wert eines ausgewählten Parameters jedes Lautes in dem Lautstrang gespeichert und der Lautstrang unter Verwendung der gespeicherten Werte in synthetische Sprache übersetzt wird, gekennzeichnet durch

Feststellen eines variablen Lautes mit einem Doppelkonsonanten-(Q), einem Silbennasal-(N) oder einen langgestreckten Laut (L) in dem Lautstrang, wobei der Wert des ausgewählten Parameters eine Funktion der Identität eines beeinflussenden Lautes in dem Lautstrang ist, und

Feststellen der Identität des beeinflussenden Lautes, der ebenfalls ein Doppelkonsonant-(Q), ein Silbennasal-(N) oder ein langgestreckter Laut (L) ist und von dem Ergebnis der Bestimmung abhängig ist, die den gespeicherten Wert des ausgewählten Parameters modifiziert, wobei der genannte beeinflussende Laut an einem vorbestimmten Ort angeordnet ist, der nichtbenachbart zu dem variablen Laut in dem Lautstrang ist.

Revendications

1. Appareil pour créer des paroles de synthèse à partir d'une suite ou d'un chapelet de caractères qui comprend:

des moyens (1) pour traduire le chapelet de caractères en une suite ou un chapelet de phonèmes, des moyens (7) pour conserver la valeur d'un paramètre sélectionné de chaque phonème contenu dans le chapelet, et des moyens de sythèse (8) pour assembler le chapelet de phonèmes en une parole ou en un

discours synthétique utilisant les valeurs conservées, caractérisé par des moyens (3) couplés auxdits moyens de traduction pour détecter les phonèmes variables présents dans le chapelet de phonèmes et, notamment, les doubles consonnes (Q), les nasales syllabiques (N) et les sons prolongés (L),

la valeur du paramètre sélectionné étant fonction de l'identité d'un phonème influent présent aussi
5 dans ledit chapelet; et

des moyens (5) couplés aux moyens de détection (3) des phonèmes variables pour déterminer l'identité du phonème influent qui est aussi soit une double consonne (Q), soit une nasale syllabique (N), soit un son prolongé (L), lesdits moyens de détermination de la valeur du paramètre répondant au résultat de la détermination en modifiant la valeur conservée du paramètre sélectionné, ledit phonème influent
10 étant situé à un emplacement prédéterminé non-adjacent à la position du phonème variable dans la suite ou le chapelet.

2. Appareil selon la revendication 1, caractérisé en ce que le paramètre sélectionné est la durée.

3. Appareil selon la revendication 1 ou 2, caractérisé en ce que les moyens de traduction comprennent une mémoire (2) pour conserver une suite ou un chapelet de phonèmes correspondant à chacun de
15 plusieurs mots, les moyens de traduction comportant des moyens de comparaison pour émettre le chapelet de phonèmes correspondant quand le mot conservé correspond à une partie de la suite ou du chapelet de caractères.

4. Appareil selon l'une quelconque des revendications 1 à 3, caractérisé en ce qu'il comprend, en outre, une mémoire (4) associée auxdits moyens de détection (3) des phonèmes variables pour conserver, en
20 association avec un phonème variable, l'emplacement relatif auquel un phonème influent va influencer la valeur du paramètre du phonème variable conservé, les moyens de détection (3) des phonèmes variables comprenant des moyens pour émettre l'emplacement prédéterminé associé quand un phonème variable est détecté dans la suite de phonèmes.

5. Appareil selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend,
25 en outre, une mémoire (6) associée aux moyens de détection (5) des phonèmes influents pour conserver, en association avec l'identité d'un phonème influent, des données représentatives d'une modification apportée à la valeur conservée du paramètre sélectionné, les moyens de détection (5) des phonèmes influents incluant des moyens de comparaison afin d'émettre les données associées quand le phonème influent conservé occupe un emplacement prédéterminé.

6. Procédé pour créer des paroles de synthèse à partir d'une suite ou d'un chapelet de caractères qui
30 consiste:

à traduire la suite ou le chapelet de caractères en un chapelet de phonèmes, à conserver la valeur d'un paramètre sélectionné de chaque phonème du chapelet et à assembler ledit chapelet de phonèmes en parole de synthèse en utilisant les valeurs conservées, caractérisé en ce que:

à détecter la présence dans la suite ou dans la chapelet de phonèmes d'un phonème variable
35 comprenant une double consonne (Q), une nasale syllabique (N) ou un son prolongé (L), la valeur du paramètre sélectionné étant fonction de l'identité d'un phonème influent présent dans le chapelet de phonèmes; et,

à déterminer l'identité du phonème influent qui est aussi soit une double consonne (Q), soit une nasale
40 syllabique (N), soit un son prolongé (L) et en réponse au résultat de la détermination, on modifie la valeur conservée du paramètre sélectionné, ledit phonème influent étant situé à un emplacement prédéterminé non-adjacent au phonème variable dans la suite ou dans le chapelet de phonèmes.

45

50

55

60

65

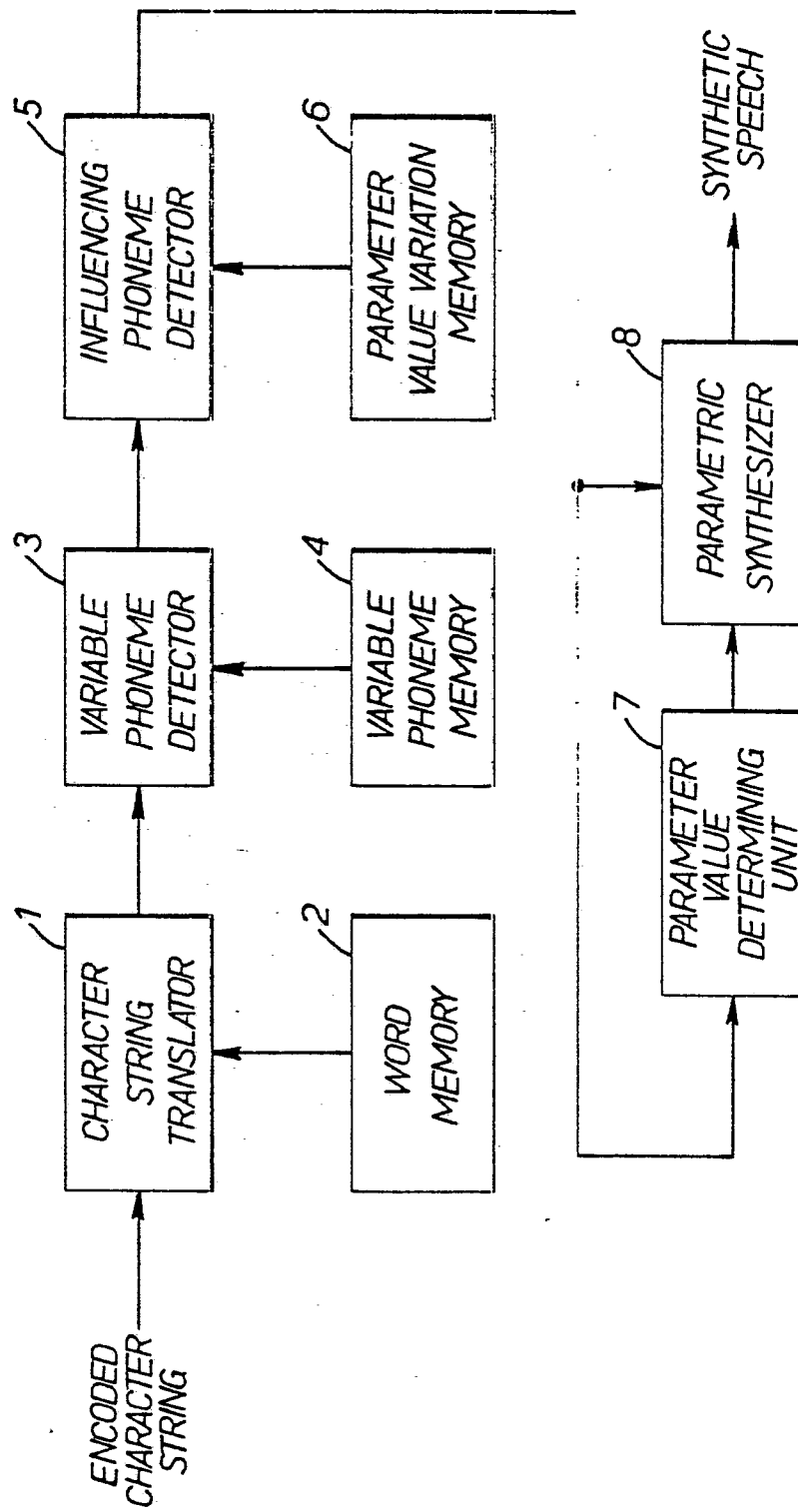
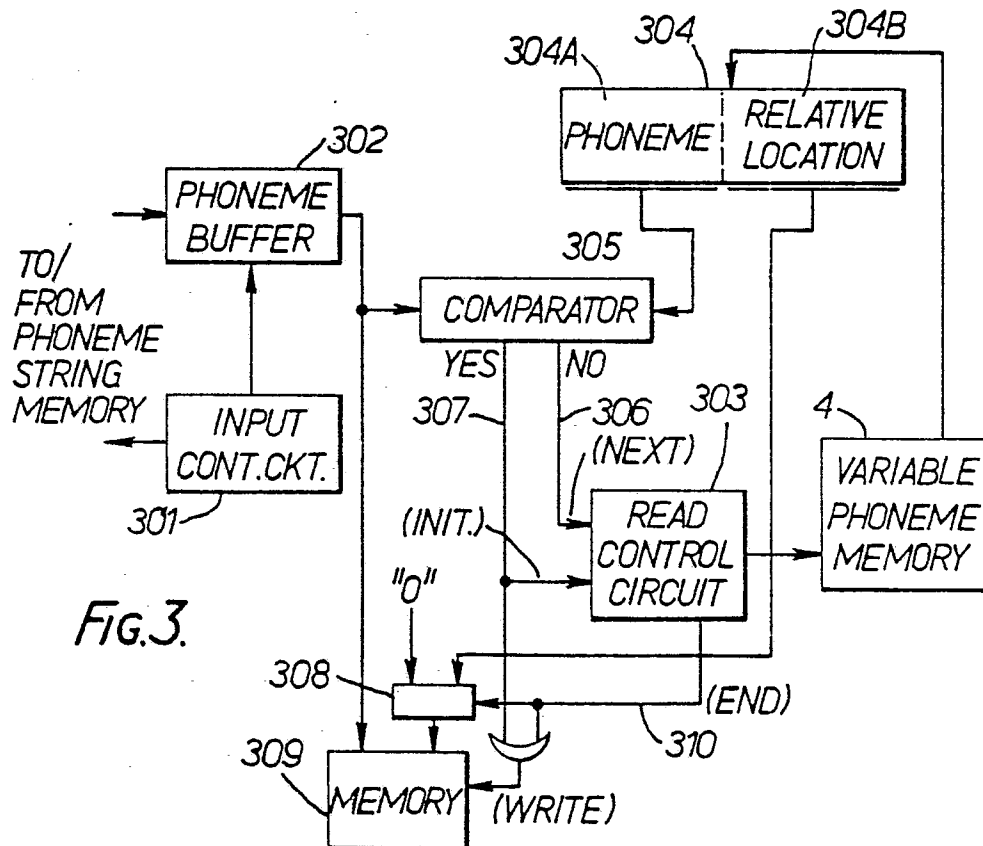
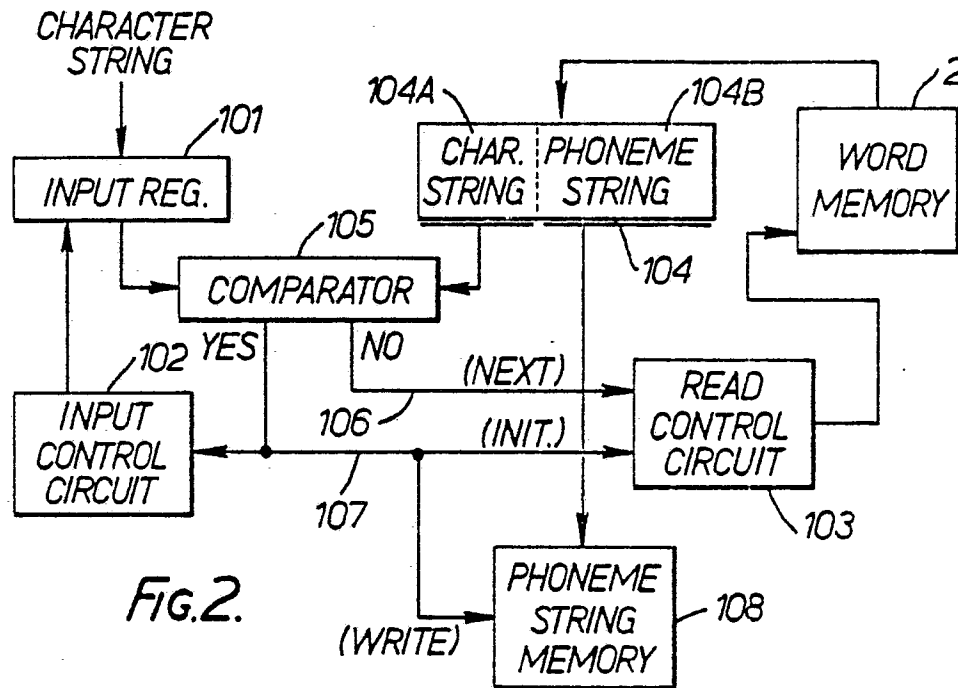


FIG. 1.



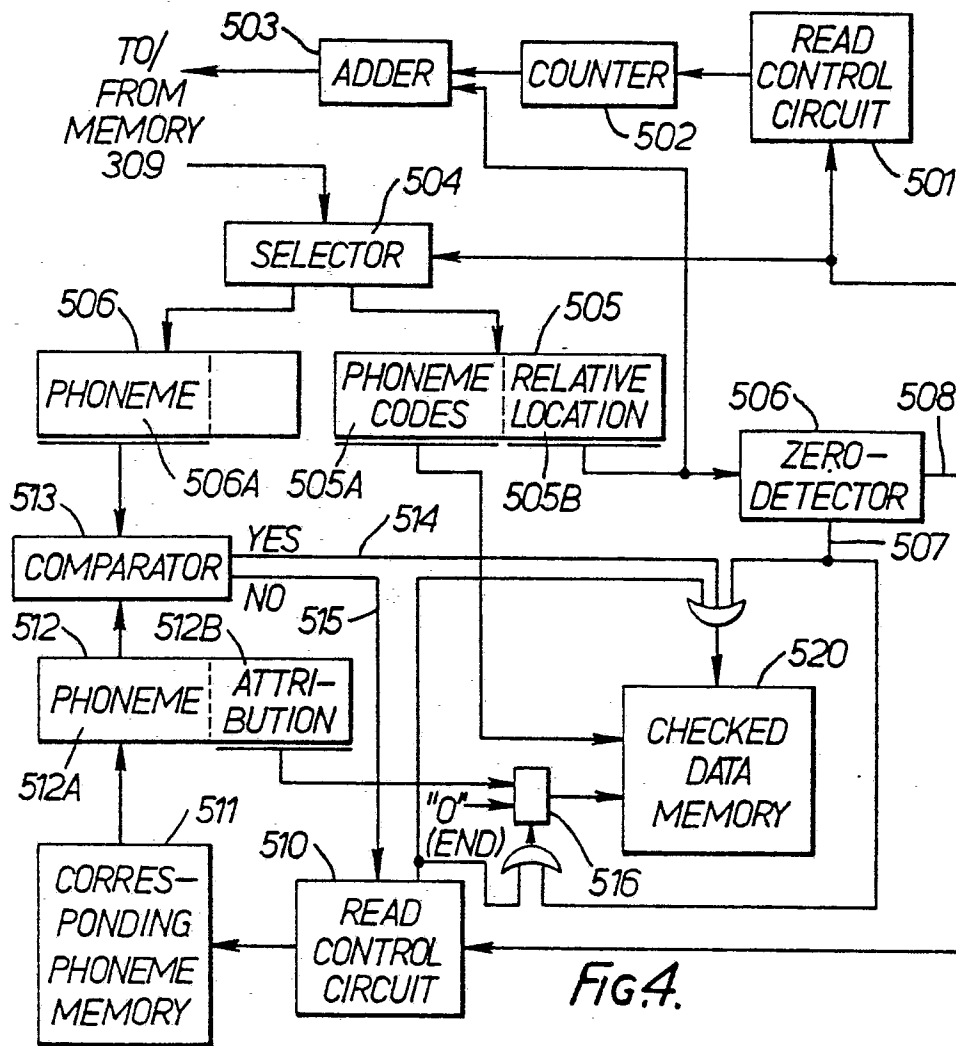


Fig. 4.

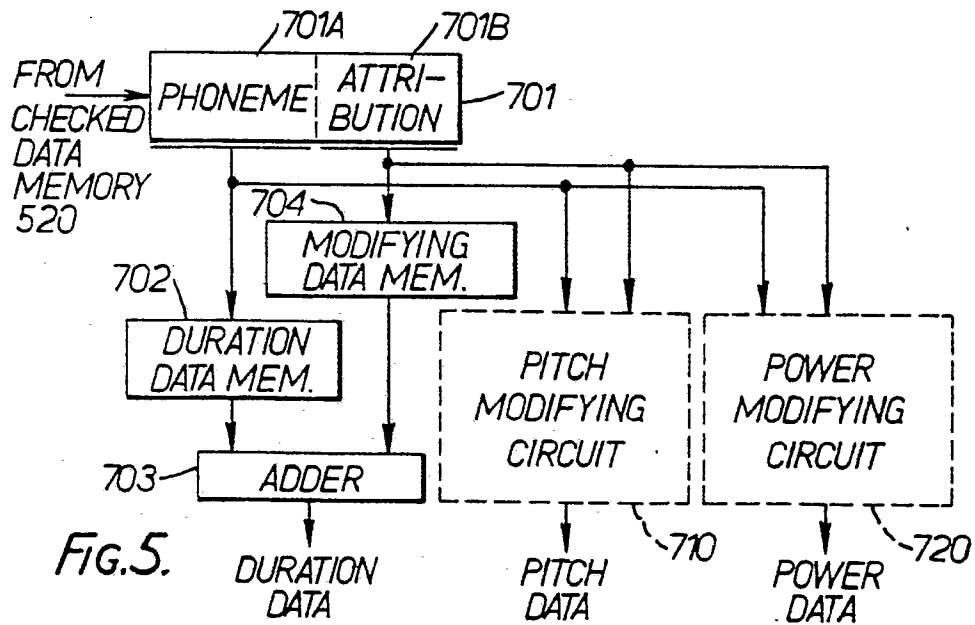


Fig. 5.

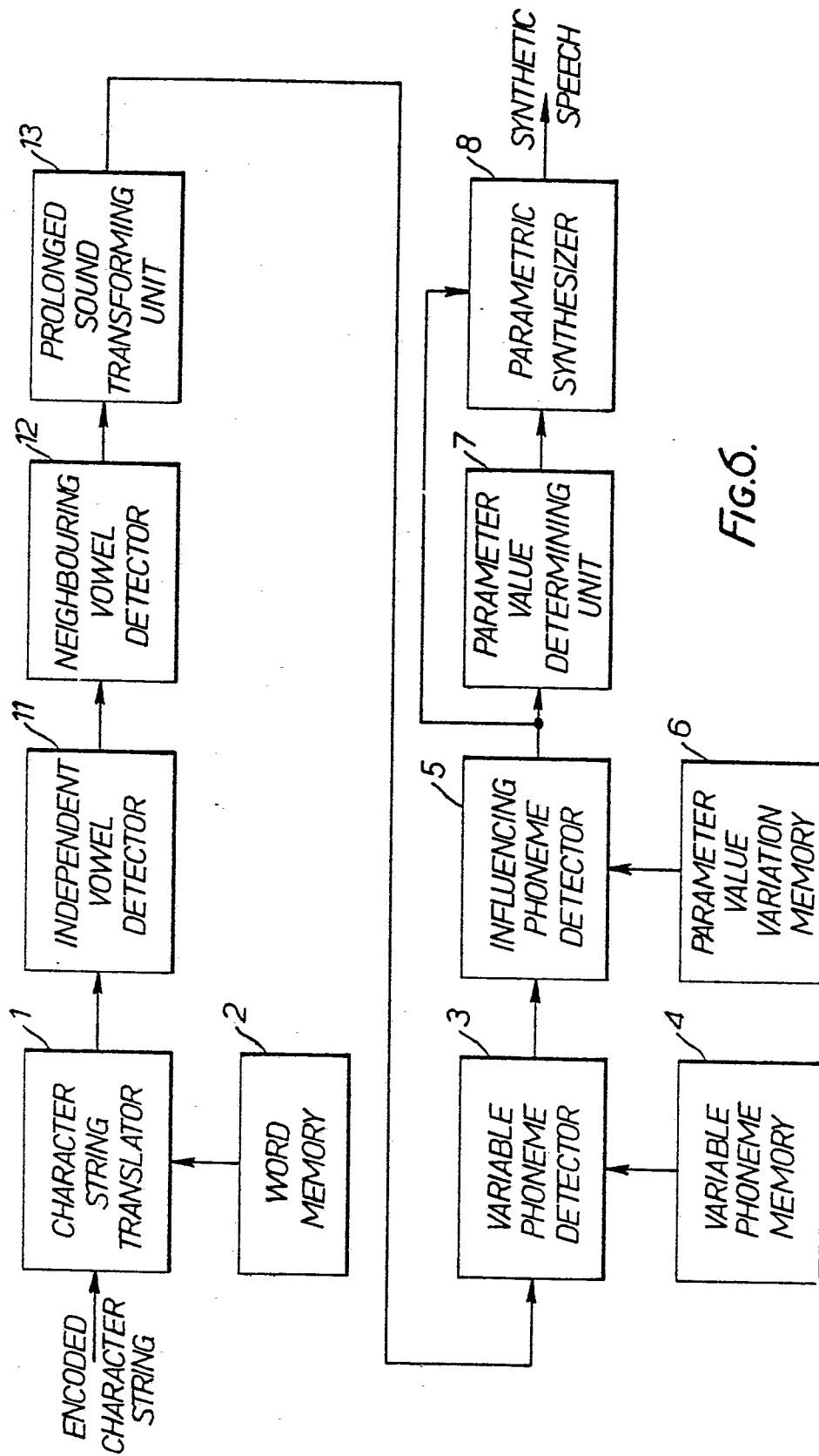


FIG. 6.