**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| (51) International Patent Classification 6 : G10L 9/06 | A1 | (11) International Publication Number: WO 99/00790 |
|---|---|---|
| | | (43) International Publication Date: 7 January 1999 (07.01.99) |

(21) International Application Number: PCT/US98/13118

(22) International Filing Date: 26 June 1998 (26.06.98)

(30) Priority Data:
60/050,998 27 June 1997 (27.06.97) US

(71) Applicant: M.H. SEGAN LIMITED PARTNERSHIP [US/US]; 126 West Avenue, Great Barrington, MA 02130 (US).

(72) Inventor: SEGAN, Marc, H.; 1192 Park Avenue, New York, NY 10128 (US).

(74) Agent: PAVANE, Martin, B.; Cohen, Pontani, Lieberman & Pavane, Suite 1210, 551 Fifth Avenue, New York, NY 10176 (US).

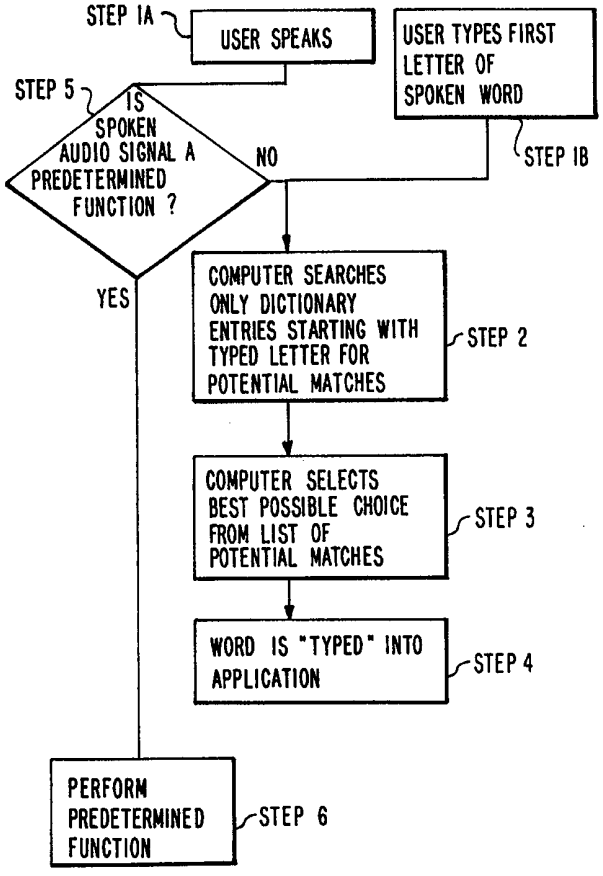
(81) Designated States: JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*

(54) Title: SPEECH RECOGNITION COMPUTER INPUT AND DEVICE

(57) Abstract

A method for recognizing a spoken audio signal as input data to a computer includes receiving the spoken audio signal at an input device of the computer (step 1A), receiving a typed first letter of the spoken audio signal at the input device of the computer (step 1B), and searching through entries that begin with the typed first letter in a dictionnary in the computer for a best match with the spoken audio signal (step 2).



*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **AL** | Albania | **ES** | Spain | **LS** | Lesotho | **SI** | Slovenia |
| **AM** | Armenia | **FI** | Finland | **LT** | Lithuania | **SK** | Slovakia |
| **AT** | Austria | **FR** | France | **LU** | Luxembourg | **SN** | Senegal |
| **AU** | Australia | **GA** | Gabon | **LV** | Latvia | **SZ** | Swaziland |
| **AZ** | Azerbaijan | **GB** | United Kingdom | **MC** | Monaco | **TD** | Chad |
| **BA** | Bosnia and Herzegovina | **GE** | Georgia | **MD** | Republic of Moldova | **TG** | Togo |
| **BB** | Barbados | **GH** | Ghana | **MG** | Madagascar | **TJ** | Tajikistan |
| **BE** | Belgium | **GN** | Guinea | **MK** | The former Yugoslav Republic of Macedonia | **TM** | Turkmenistan |
| **BF** | Burkina Faso | **GR** | Greece | | | **TR** | Turkey |
| **BG** | Bulgaria | **HU** | Hungary | **ML** | Mali | **TT** | Trinidad and Tobago |
| **BJ** | Benin | **IE** | Ireland | **MN** | Mongolia | **UA** | Ukraine |
| **BR** | Brazil | **IL** | Israel | **MR** | Mauritania | **UG** | Uganda |
| **BY** | Belarus | **IS** | Iceland | **MW** | Malawi | **US** | United States of America |
| **CA** | Canada | **IT** | Italy | **MX** | Mexico | **UZ** | Uzbekistan |
| **CF** | Central African Republic | **JP** | Japan | **NE** | Niger | **VN** | Viet Nam |
| **CG** | Congo | **KE** | Kenya | **NL** | Netherlands | **YU** | Yugoslavia |
| **CH** | Switzerland | **KG** | Kyrgyzstan | **NO** | Norway | **ZW** | Zimbabwe |
| **CI** | Côte d'Ivoire | **KP** | Democratic People's Republic of Korea | **NZ** | New Zealand | | |
| **CM** | Cameroon | | | **PL** | Poland | | |
| **CN** | China | **KR** | Republic of Korea | **PT** | Portugal | | |
| **CU** | Cuba | **KZ** | Kazakstan | **RO** | Romania | | |
| **CZ** | Czech Republic | **LC** | Saint Lucia | **RU** | Russian Federation | | |
| **DE** | Germany | **LI** | Liechtenstein | **SD** | Sudan | | |
| **DK** | Denmark | **LK** | Sri Lanka | **SE** | Sweden | | |
| **EE** | Estonia | **LR** | Liberia | **SG** | Singapore | | |

# SPEECH RECOGNITION COMPUTER INPUT AND DEVICE

## BACKGROUND OF THE INVENTION

### 1. Field of the Invention

The present invention relates to a computer interface method and device wherein a user's voice is used to input data to a computer. More particularly, the present invention pertains to a method and apparatus wherein a user's voice, in conjunction with a keyboard, is used for inputting instructions or data to a computer. Most particularly, the present invention relates to a voice recognition computer input method and apparatus wherein a user inputs data into a computer by speaking each word while typing, via a keyboard, the first letter of each spoken word to assist the computer in identifying or recognizing the spoken words.

### 2. Description of the Related Art

There currently exists various inexpensive software programs for use with personal computers that allow voice input of text and instructions. Some such programs are offered by Dragon Systems, IBM and Kurzweil. These programs allow a user's speech to be input to a computer via, for example, an audio input device, whereupon the spoken words are identified and used for various applications such as for text input or to execute commands. As can be appreciated, such man-machine dictation programs save users time by alleviating the need to enter data via a keyboard to generate documents.

With reference to FIG. 1, a diagram of the steps used in a prior art program for effectuating speech input to a computer is depicted. As shown,

the prior art program is divided into five separate steps. The first step involves the user speaking in proximity to a computer so that a sound detector can receive the spoken audio signal. The computer then determines when each word begins. For each spoken word, a dictionary search is conducted to locate potential matches between stored dictionary words and each spoken word. The best dictionary word match is then selected from the potential matches and the selected word is forwarded to an application.

Unfortunately, such prior art programs suffer several drawbacks. Specifically, the error rate for such programs remains high. Moreover, problems exist with continuous speech input, thus requiring an awkward style of spoken input by pausing at intervals to allow the program to identify the spoken words and process the information. Also, such prior art programs must be trained prior to use. In other words, prior to use a user must take time to record a large sample vocabulary for the machine to compare with inputs.

As a result of the drawbacks of prior art programs, users are left with the situation of continuing to rely upon old methods for accurate, convenient text input, which involve either someone to take dictation and then enter the spoken words into a computer via a keyboard, or fast typing skills. Moreover, the continuous and frequent use of a keyboard results in a risk of repetitive stress injuries and fatigue.

## BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings:

FIG. 1 is a block diagram depicting the steps of a prior art voice input method;

FIG. 2 is a block diagram depicting the steps of an embodiment of the voice recognition computer input method and device in accordance with the present invention; and

FIG. 3 is a block diagram depicting the steps of another embodiment of the voice recognition computer input method and device in accordance with the present invention

## DETAILED DESCRIPTION OF THE PRESENTLY PREFERRED EMBODIMENTS

The present invention utilizes known voice recognition technology to allow a user to type the initial letter of each spoken word, via a keyboard, substantially simultaneously with the spoken word, thereby reducing the processing power and time required by a computer. The invention overcomes the drawbacks of the prior art in several respects. Initially, continuous speech becomes possible because the computer is advised exactly when a new word begins -- by keyboard entry of the first letter of each word -- thus, vastly reducing processing power which was required in prior art systems to determine when one word ended and another began. In addition, the possible choice of words is vastly reduced for the computer since the initial letter is known via keyboard input. This feature further reduces the required processing power and time because fewer words from the dictionary need to be located and retrieved. This feature also increases the accuracy of the selected or retrieved word

because the first letter of the spoken word is provided to the computer.

In addition to the foregoing benefits, the present invention requires significantly less typing than standard keyboard entry. As a result, smaller keyboards relative to standard sized keyboards may be used, thereby increasing user workstation surface area that was previously occupied by a standard size keyboard. Moreover, the present invention is particularly well-suited for use with palm-sized computers or Personal Digital Assistants (PDA's). Furthermore, as less typing is required for data input, there is a reduced chance that injury, fatigue and error will occur.

As a result of the present invention, speaker independent programs can become feasible much more easily since the reliance on the computer's ability to discriminate variations in pronunciation is considerably diminished.

The present invention is depicted in Fig. 2. As shown, steps 1A and 1B are performed substantially simultaneously with each other. In the alternative, a buffer register may be employed to temporarily store either the spoken word or the typed letter while awaiting input of the other of the corresponding spoken word or input letter. Once both a spoken word and a letter are input to the computer, a search is conducted among dictionary entries, starting with the typed letter, to locate potential matches. The best match is then selected and the selected word is forwarded to an application.

In addition to the above, Fig. 3 shows that the present invention may assign or designate

single or multiple words to specific functions which are performed when the designated word is spoken without the first letter of the word being typed as in steps 5 and 6 of Fig. 3. As an illustration, the word "delete" may be assigned the function of erasing the last typed word when the word "delete" is spoken without typing the letter "d". Thus, if the computer incorrectly identifies a word based on the spoken word and typed in first letter of the spoken word, a user can strike or erase the incorrect word by saying the word "delete". Of course, if the word "delete" is intended to be input, the word will be spoken and the letter "d" will be pressed on the keyboard. Such a function can be performed on any incorrect word by simply moving the cursor to the incorrect word and saying "delete". Step 5 comprises the step of comparing the spoken audio signal to a list of predetermined functions. Step 5 may optionally comprise the steps of determining whether the user types a first letter and then comparing the spoken audio signal to the list of predetermined functions.

The invention is not limited by the embodiments described above which are presented as examples only but can be modified in various ways as will be readily recognized by those having ordinary skill in the art. For example, instead of typing only the first letter of each spoken word, the first and last letters can be typed, or, alternatively, the first two letters can be typed, thereby further reducing computer processing power and time.

**CLAIMS**

1. A method for facilitating the identification of a spoken audio signal in a voice recognition computer application, the method comprising the steps of:

a. receiving a spoken audio signal at an input device of a computer;

b. receiving a typed first letter of the spoken audio signal at the input device of the computer;

c. after steps a. and b., searching entries starting with the typed first letter in a dictionary of the computer and selecting a best match; and

d. inserting the best match into a currently running application of the computer.

2. The method of claim 1, wherein said steps a. and b. are performed substantially simultaneously.

3. The method of claim 1, wherein one of said steps a. and b. is performed before the other of said steps a. and b., and said computer interface method further comprises the step of storing said one of said steps a. and b. in a buffer register until the other of said steps a. and b. is performed.

4. The method of claim 1, further comprising the steps of:

f. determining whether the spoken audio signal comprises a predetermined function;

g. performing the predetermined function if it is determined in step f. that the spoken audio signal comprises the predetermined function; and

h. performing step c. if the spoken audio interface is not the predetermined function.

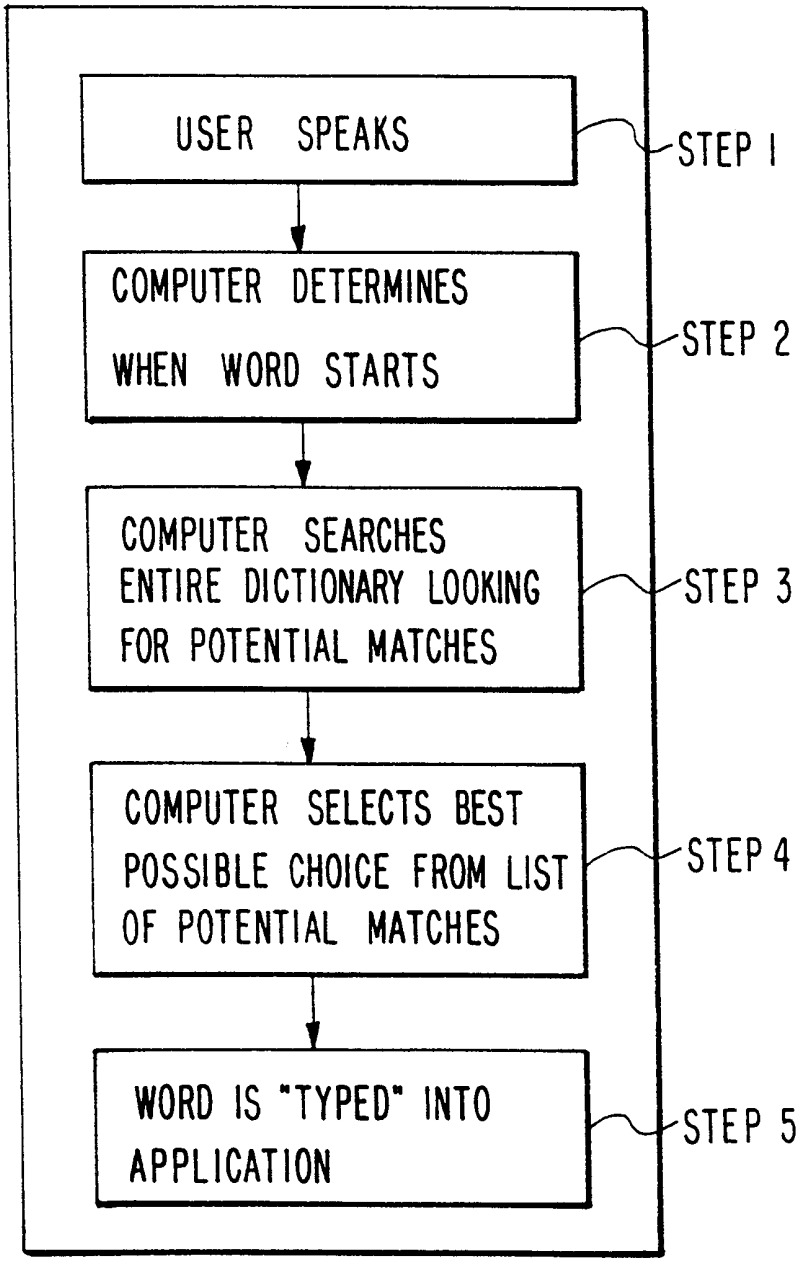



*FIG. 1*
PRIOR ART

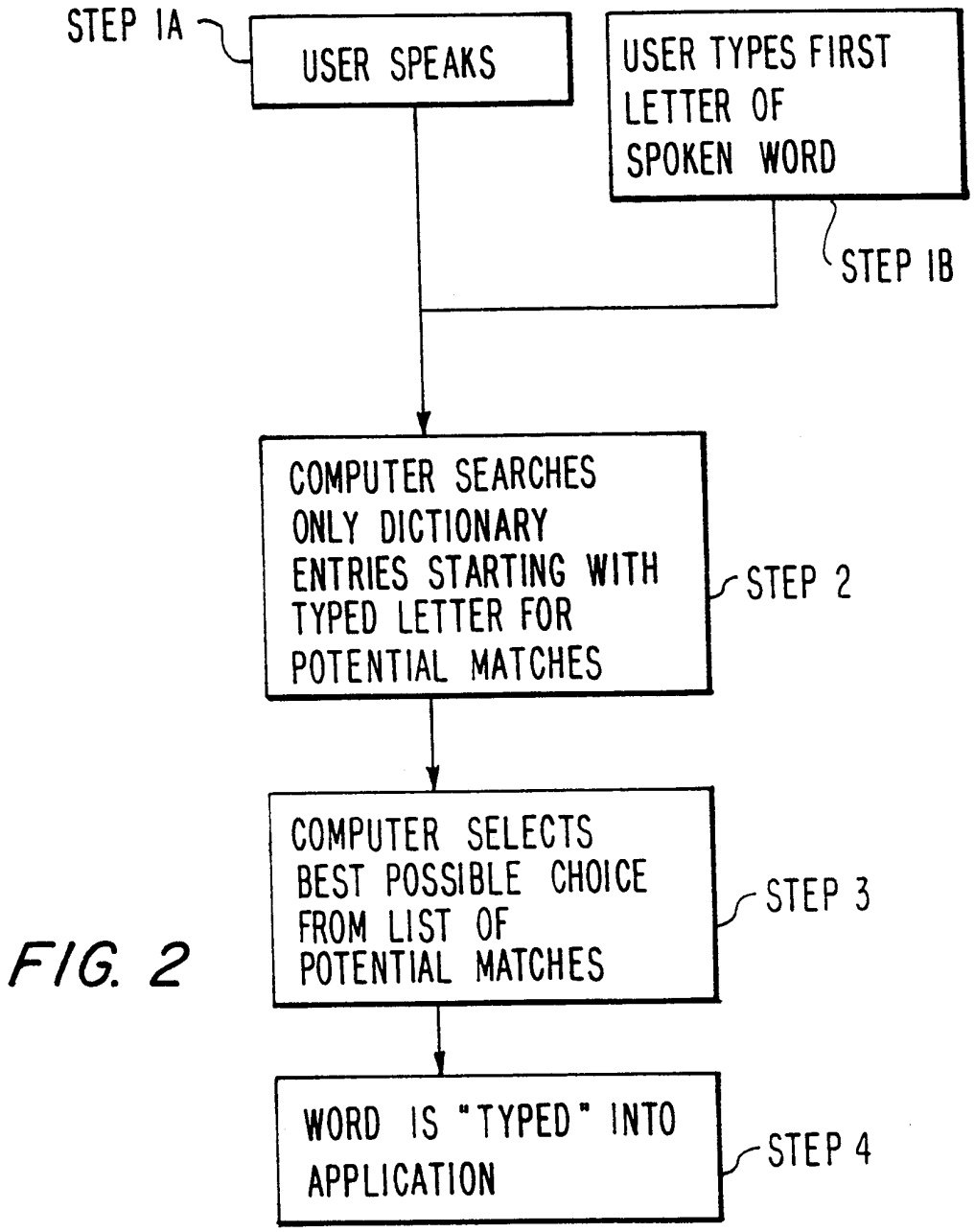

STEP 1A
USER SPEAKS
USER TYPES FIRST LETTER OF SPOKEN WORD
STEP 1B
COMPUTER SEARCHES ONLY DICTIONARY ENTRIES STARTING WITH TYPED LETTER FOR POTENTIAL MATCHES
STEP 2
COMPUTER SELECTS BEST POSSIBLE CHOICE FROM LIST OF POTENTIAL MATCHES
STEP 3
WORD IS "TYPED" INTO APPLICATION
STEP 4

FIG. 2

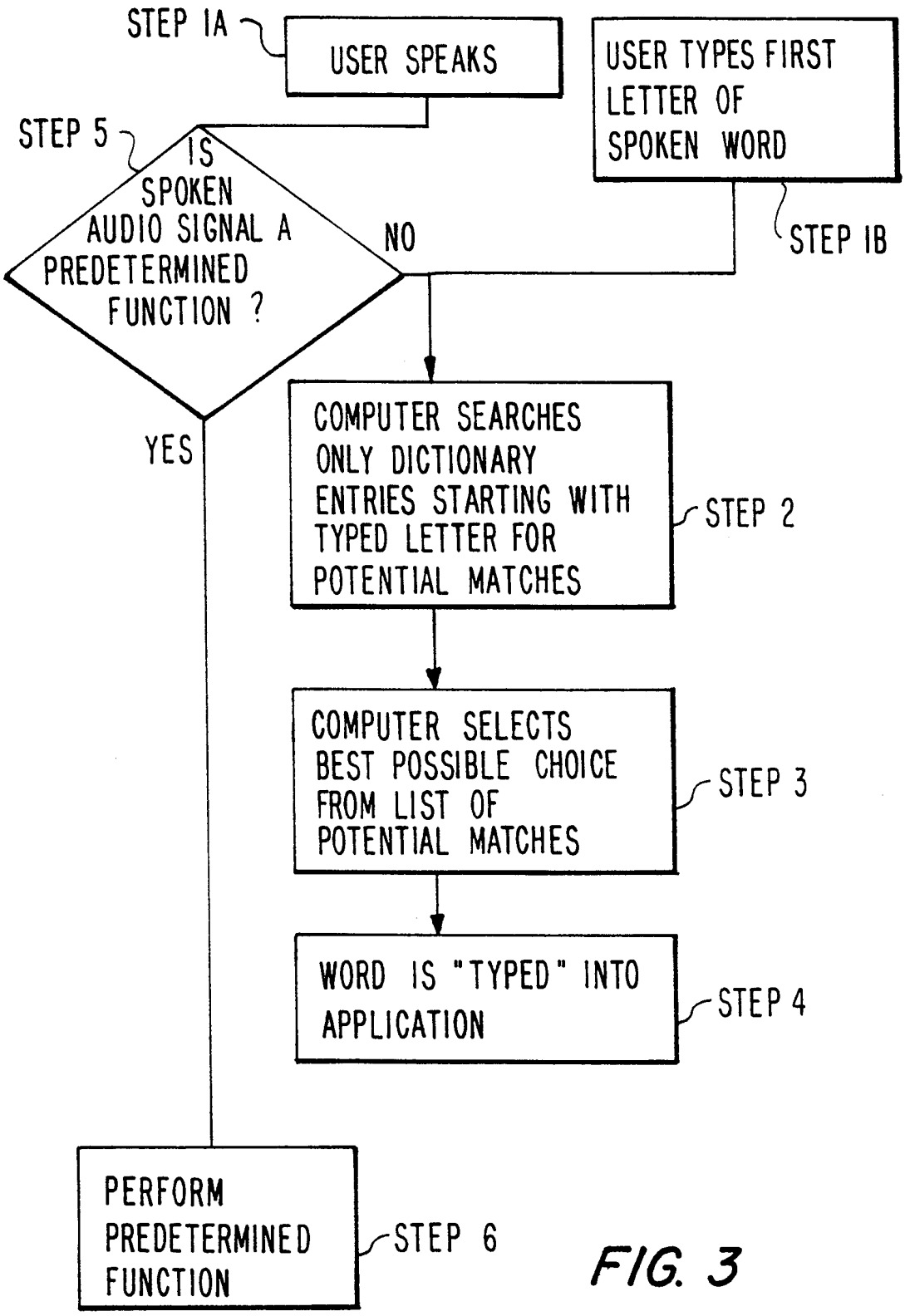

STEP 1A
USER SPEAKS
USER TYPES FIRST LETTER OF SPOKEN WORD
STEP 1B
STEP 5
IS SPOKEN AUDIO SIGNAL A PREDETERMINED FUNCTION ?
NO
YES
COMPUTER SEARCHES ONLY DICTIONARY ENTRIES STARTING WITH TYPED LETTER FOR POTENTIAL MATCHES
STEP 2
COMPUTER SELECTS BEST POSSIBLE CHOICE FROM LIST OF POTENTIAL MATCHES
STEP 3
WORD IS "TYPED" INTO APPLICATION
STEP 4
PERFORM PREDETERMINED FUNCTION
STEP 6

FIG. 3

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/13118

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :G10L 9/06
US CL :704/235, 275

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/235, 252, 254, 275

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

SmartPatents

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | US 5,027,406 A (ROBERTS et al) 25 June 1991, Abstract, col. 14, lines 19-21;<br>col 8, lines 17-20 and col 25, lines 6-9;<br>col. 8, lines 30-32 and 37-39, Abstract | 1,2<br>3<br>4 |
| Y<br>---<br>A | US 4,866,778 (BAKER) 12 September 1989, Abstract and Figures 3 and 7 with accompanying discussion. | 1-3<br>---<br>4 |
| A | US 5,022,081 (HIROSE, et al) 4 June 1991, passim. | 1-4 |
| A | US 5,329,609 (SANADA, et al) 12 July 1994, passim. | 1-4 |

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

| | | | |
|---|---|---|---|
| * | Special categories of cited documents: | "T" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| "A" | document defining the general state of the art which is not considered to be of particular relevance | "X" | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| "E" | earlier document published on or after the international filing date | "Y" | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "L" | document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "&" | document member of the same patent family |
| "O" | document referring to an oral disclosure, use, exhibition or other means | | |
| "P" | document published prior to the international filing date but later than the priority date claimed | | |

| | |
|---|---|
| Date of the actual completion of the international search<br>15 SEPTEMBER 1998 | Date of mailing of the international search report<br>19 OCT 1998 |
| Name and mailing address of the ISA/US<br>Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231<br>Facsimile No. (703) 305-9508 | Authorized officer<br>TALIVALDIS IVARS SMITS<br>Telephone No. (703) 306-3011 |

Form PCT/ISA/210 (second sheet)(July 1992)★