

SPECIFICATION

Method and apparatus for the composition of texts in arabic characters

The invention relates to apparatus and a method for the composition of texts in Arabic characters.

The written, typed or printed Arabic language has different shapes for certain characters, so-called basic characters, which depend on the characteristics of the preceding and following characters. In fact, a same character is not shaped identically if it is independent, located at the beginning or at the end of a word or further if it is linked to the preceding character.

In order to cater for these different shapes, the existing keyboards have keys for each different shape of the same character. In the example shown in Figures 1 of the accompanying drawings, the different shapes of the basic character "ain" appear on keys 8, 19 and 39 of the keyboard. This leads, in order to maintain the number of keys to an acceptable value, to use a same key for two different shapes which may correspond to different basic characters and necessitates the use of a special key to select the print of the desired shape. This leads to complicated manipulations causing frequent typing errors.

The operator must in fact first select the key on which appears the desired shape of a given basic character then select this shape from the other sign appearing on said same key by means of a special key 35 or 46.

According to the present invention in one aspect there is provided a method for the composition of texts in Arabic characters, in which at least certain basic letters which can have four different shapes correspond each to only one key of the keyboard; that when using said key one determines the shape to be used of the basic letters according to the characteristics of the keys used just after, or just before and after the said key, the preceding key used to select the shape of the desired letter being a key corresponding either to a letter that is linked with a following letter or to any other letter or sign or function which is not linked with a following letter, and the following key used to select the shape of the desired letter being a key which represents either any letter whether linked or not or any other sign or function which is not a letter.

According to the present invention in another aspect there is provided apparatus for the composition of texts in Arabic characters, comprising a keyboard having several keys each corresponding to only one basic character in all its four different shapes; an electronic device having a memory retaining the codes corresponding to all the shapes of each basic character and a working memory retaining the characteristics of the keys depressed immediately before and after the considered key; and a logic circuit for determining, according to the characteristics memorized in the working memory, the codes corresponding to the appropriate shape of the basic character, whereby said apparatus reproduces only that shape of a basic character that corresponds to the keys depressed immediately before and after the key corresponding to the basic character.

An embodiment of the present invention will now be described, by way of an example, with reference to the accompanying drawings, in which:—

Figure 1 shows a conventional typewriter keyboard,

Figure 2 shows a typewriter keyboard according to the present invention,

Figure 3 shows a telex keyboard according to the present invention having Arabic as well as Latin characters,

Figure 4 shows a block diagram of the apparatus,

Figure 5 shows the four shapes of the basic character "ain",

Figure 6a is an example of a character linked to the following character, and

Figure 6b is an example of a character which is separated from the following character.

According to the method of the present invention, each basic character corresponds to only one key of the keyboard. For the basic character having in the composition of texts several shapes, only one shape, for example the primary shape, appears on the key corresponding to said basic character.

To facilitate the use of the keyboard, the key corresponding to a given basic character is located at the same location of the keyboard as the key corresponding to the first shape of this character in a conventional keyboard.

When depressing a key corresponding to a basic character, the selection of the appropriated shape of this basic character is made according to the keys depressed before, or before and after, the considered key.

The selected shape of the basic character depends on the one hand to the fact that the key depressed afterwards corresponds either to a letter whether linked or not (not N) or to a sign or function which is not a letter (N). This shape may also depend from the fact that the key depressed before corresponds either to a linked letter (L) or to another sign, letter or function which is not a linked letter (not L).

In the case of a typewriter, the shape of certain basic letters is determined by the two following keys. This case is present for the basic letters appearing on keys 15, 26, 27, 30, 31 and 36 when the first key used after is one of the keys 26 or 36 or by the characteristics of the first following keys when said key is one of the keys 37, 39 or 44 (Figure 2).

The keys of the keyboard shown at Figure 2 may be divided up into three categories according to their characteristics. These three categories are:

1. The linked letters L (see Figure 6a).

These letters are the ones which are linked to the following letter in writing.

In the keyboard shown at Figure 2 these letters correspond to keys 3, 4; 13 to 28; 30 to 33 and 36.

2. The independent (or unlinked) letters D (see Figure 6b).

These independent letters are the ones which are not linked to the following letter in writing. They correspond in the keyboard shown at Figure 2 to the keys 5, 6, 29, 37, 39, 40, 43 and 44.

3. The keys which do not correspond to a letter (N).

This category of keys includes the one corresponding to the signs which are the following keys of keyboard of Figure 2; in letters and figures positions 1, 2, 7, 8, 9, 10, 11, 12, 34, 42, and in figures positions 3, 4, 5, 6 as well as the two keys 38 and 45.

Furthermore this category of keys includes all the keys corresponding to functions of the machine such as advance paper, carriage movement, etc.

The following table shows the selection of one of the four shapes of the basic character "ain" (see Figure 5) according to the characteristics of the preceeding and following keys used.

Preceding Key	Following Key	Selected shape (see Figure 5)
D	D	1
D	L	1
not L	not N	
N	D	1
N	L	1
L	D	2
L	not N	
L	L	2
L	N	3
D	N	4
not L	N	
N	N	4

For each basic character having several shapes, the electronic part of the keyboard comprises a logic similar to the one given by way of example for the character "ain" selecting the desired shape as a function of the key depressed after, or before and after, the one corresponding to the desired key.

The codes which are created by the described keyboard can be used, possibly after decoding, to control any kind of printing and/or video machines such as a typewriter, a composing machine, a telex machine, etc.

The device used for the realization of this method of composition comprises a keyboard having keys corresponding to the different signs having to be transmitted or printed. The keyboard can be made as shown for example in Figures 2 or 3.

In this keyboard the basic characters which may present several shapes according to their use correspond to only one key. This keyboard controls an electronic device which in practice utilizes micro-processors but one example of which is schematically given here in the form of a block diagram at Figure 4.

In this Figure 4, a represents the keyboard; b a control device (micro-processor LSI); c a keyboard encoder (LSI); d a memory (read only memory LSI); e a working flag register (read after memory LSI); f

an input/output port (LSI) and *g* an output device.

The functions of these different elements of the electronic device are the following:

b — detects the informations of the keyboard code memorized in *f*.

5 — refers to the memory *d* containing the list of the output codes corresponding to all shapes of all letters or signs having to be printed or transmitted 5

— selects one or more output codes from *d* and transmit them

— causes the modification of the informations memorized in *e*.

The working sequences of the control device *b* are effected according to a program which is also recorded in the memory *d*.

10 In the following the different steps made by this device are given for the selection of one shape of 10 a basic character by means of the depression of the one key of the keyboard corresponding to all possible shapes of said basic character.

One admits in the following that the working flag register *e* comprises at the beginning of the sequence of operations the following informations:

15 — in its first part: the characteristic (L or not L) of a key which has just been depressed before. 15

— in its second part: the keyboard code corresponding to the key depressed immediately before, provided that this key corresponds to a letter having more than one shape. If not this second part is empty.

(1) A key of the keyboard is depressed.

20 (2) The depression of said key causes the formation of a keyboard code corresponding to said key 20 which is obtained by the keyboard encoder *c*.

(3) The keyboard encoder *c* sends the keyboard code corresponding to the depressed key to the input/output port *f* which stores it temporarily.

25 (4) The control device *b* detects some informations which are present in port *f*. In fact the keyboard 25 code of the depressed key permits to know if this key corresponds to a character having only one shape or a character having several shapes.

(5) The control device *b* detects in the first part of the working flag register *e* if the second part of this register stores a keyboard code corresponding to a basic character having several shapes, the output code of the shape to be printed having not yet been completely selected. One admits in the 30 present case that the second part of the flag register *e* is empty. 30

(6) The control device *b* detects in the first part of the working flag register *e* if the last printed character was a linked (L) letter or an unlinked (not L) letter, sign or function.

(7) The control *b* looks in the memory *d* and selects all the output codes corresponding to the different shapes which can be printed referring to the depressed key. In the following the 35 depressed key is taken to be the one corresponding to the letter "ain" (Figure 5) which has four 35 possible shapes.

(8) Depending on the information present in the first part of the working flag register *e* (L or not L) *b* cancels two of the four output codes of the letter "ain". In the case where the characteristic of the preceding key was L (linked letter) the control *b* cancels the output codes corresponding 40 to the shapes 1 and 4 of Figure 5. In the opposite case (unlinked letter not L) the shapes 2 and 3 40 of Figure 5 are cancelled (see table).

(9) The control device *b* erases the flag register *e* and delivers new information to said register.

a) In the first part of the flag register if the "ain" letter is linked (L) or not (not L), as well as if it is definitely selected or not. In this case the selection is not completed.

45 (b) In the second part of the flag register *e* the keyboard code of the letter "ain" is recorded by 45 transmission from port *f* to the register *e*. This second part of the flag register receives this information only if the keyboard code corresponds to more than one output code and this is the case for the letter "ain".

50 (10) The next key is depressed and the port *f* receives and stores temporarily a keyboard code 50 generated by encoder *c*.

(11) The control device *b* detects the characteristic N or not N of the code stores in the port *f*.

(12) The control device *b* detects in the first part of the flag register *e* the presence or the absence of a keyboard code stored in the second part of this register *e*. In this case there is such a code in this 55 second part of the register *e*. The control *b* also knows that two output codes have already been 55 cancelled.

(13) In relation with the information obtained during operation steps No. 11 the control device *b* selects in the memory *d* one of the two remaining output codes. If the keyboard code of the port *f* is N, the control device *b* cancels the output codes corresponding to the shape 1 or 2 of Figure 5, 60 whereas if the keyboard code of the port *f* is "not N" *b* cancels the output codes corresponding 60 to the shapes 3 or 4 of Figure 5 (see table). Taking into account the first selection made at steps No. 8 above and this second selection only one of the four output codes of the letter "ain" is selected.

(14) The control device *b* causes the transmission, through the memory *d*, of the unique output code selected to the output device *g*.

65 In the case where the depressed key corresponds to a basic character having three different 65

shapes, the same operations as above are effectuated.

If the depressed key corresponds to a basic character having only two different shapes corresponding either to the beginning or to the end of a word, the control device *b* knows it from step No. 7 above and step No. 8 is then cancelled.

In the case where the depressed key corresponds to a letter or sign having only one shape the selection operations Nos. 8 and 13 are cancelled.

Thanks to the characteristic of the following key (N or not N) and possibly of the preceding key (L or not L) the appropriate shape of a basic letter having to be printed is selected, and still there is only one key for all the shapes of this basic letter.

In the case where one uses this method to send messages by telex through the existing connections and exchanges, it is preferable to have a bi-alphabetical keyboard. Such key may correspond according to the working mode to an Arabic character or a Latin character.

To establish a connection or call, the operator uses for calling the receiver the Latin mode. The service signals such as busy line, out of order, etc., are automatically printed in the two alphabets and the two languages. When the connection is established the operator chooses the transmission mode, Latin or Arabic, according to the message to be transmitted.

The receiver has to be able to identify the language of the received codes and commute to the corresponding Latin or Arabic mode.

Therefore, the codes received are stored in a memory the capacity of which is greater than one line of text. When this memory is full and there is no reception of the "change of line" code in the Arabic alphabet which correspond to a sequence of codes not used in the transmission of Latin characters, the machine starts to print the received codes in the Latin mode. Conversely at the reception of this "change of line" code in Arabic mode, the machine is automatically switched over to the Arabic mode.

In the Arabic language countries almost all messages start with (the BASMALLAH): BISM ILLAH ALROUHMAN AL RAHIM in which any typing error is considered as a serious error, it is important to have the transcription of this sentence which is taken from the Coran, automatically. According to the present method there is a special key on the keyboard which permits the automatic printing of this whole sentence.

It is evident that for reasons of convenience or of transmission the electronic part (*b, c, d, e, f, g*) of the composition device can be coupled either with a printer or a keyboard.

CLAIMS

1. A method for the composition of texts in Arabic characters, in which at least certain basic letters which can have four different shapes correspond each to only one key of the keyboard; that when using said key one determines the shape to be used of the basic letters according to the characteristics of the keys used just after, or just before and after the said key, the preceding key used to select the shape of the desired letter being a key corresponding either to a letter than is linked with a following letter or to any other letter or sign or function which is not linked with a following letter, and the following key used to select the shape of the desired letter being a key which represents either any letter whether linked or not or any other sign or function which is not a letter.

2. The method claimed in claim 1 utilizing a keyboard as shown in Figure 2, in which a particular shape to be used for basic letters corresponding to keys 15, 26, 27, 30, 31 and 36 is determined by the characteristics of the two following keys when the first following key is one of the keys 26, 36 and by the characteristics of the first of said following key when the first key is 37, 39 or 44.

3. The method claimed in claim 1 or claim 2, in which the composition of the sentence BISM ILLAH ALROUHMAN AL RAHIM is automatically obtained by the use of only one particular key.

4. The method claimed in any preceding claim, when used for the transmission of a telex message using an existing network, in which one determines at reception if the signal received corresponds to a text in Arabic or not and that one switches accordingly the working mode of the receiver.

5. The method claimed in claim 4, in which this determination is made by storing the signals corresponding to at least one line of text and determining the presence or the absence in the stored signals of the sequence of codes corresponding to the function "changes of line" in the CCITT No. 2 Arabic alphabet which corresponds to a sequence of codes not used in the transmission of Latin characters.

6. The method claimed in claim 4, in which one prints automatically the service signals received from the exchanges with their translation in Arabic letters.

7. A method for the composition of texts in Arabic characters, substantially as hereinbefore described with reference to Figures 2 to 6 of the accompanying drawings.

8. Apparatus for the composition of texts in Arabic characters, comprising a keyboard having several keys each corresponding to only one basic character in all its four different shapes; an electronic device having a memory retaining the codes corresponding to all the shapes of each basic character and a working memory retaining the characteristics of the keys depressed immediately before and after the considered key; and a logic circuit for determining, according to the characteristics memorized in the working memory, the codes corresponding to the appropriate shape of the basic character, whereby said apparatus reproduces only that shape of a basic character that corresponds to the keys depressed



shapes, the same operations as above are effectuated.

If the depressed key corresponds to a basic character having only two different shapes corresponding either to the beginning or to the end of a word, the control device *b* knows it from step No. 7 above and step No. 8 is then cancelled.

5 In the case where the depressed key corresponds to a letter or sign having only one shape the selection operations Nos. 8 and 13 are cancelled. 5

Thanks to the characteristic of the following key (N or not N) and possibly of the preceding key (L or not L) the appropriate shape of a basic letter having to be printed is selected, and still there is only one key for all the shapes of this basic letter.

10 In the case where one uses this method to send messages by telex through the existing connections and exchanges, it is preferable to have a bi-alphabetical keyboard. Such key may correspond according to the working mode to an Arabic character or a Latin character. 10

To establish a connection or call, the operator uses for calling the receiver the Latin mode. The service signals such as busy line, out of order, etc., are automatically printed in the two alphabets and the two languages. When the connection is established the operator chooses the transmission mode, Latin or Arabic, according to the message to be transmitted. 15

The receiver has to be able to identify the language of the received codes and commute to the corresponding Latin or Arabic mode.

20 Therefore, the codes received are stored in a memory the capacity of which is greater than one line of text. When this memory is full and there is no reception of the "change of line" code in the Arabic alphabet which correspond to a sequence of codes not used in the transmission of Latin characters, the machine starts to print the received codes in the Latin mode. Conversely at the reception of this "change of line" code in Arabic mode, the machine is automatically switched over to the Arabic mode. 20

25 In the Arabic language countries almost all messages start with (the BASMALLAH): BISM ILLAH ALROUHMAL AL RAHIM in which any typing error is considered as a serious error, it is important to have the transcription of this sentence which is taken from the Coran, automatically. According to the present method there is a special key on the keyboard which permits the automatic printing of this whole sentence. 25

30 It is evident that for reasons of convenience or of transmission the electronic part (*b, c, d, e, f, g*) of the composition device can be coupled either with a printer or a keyboard. 30

CLAIMS

1. A method for the composition of texts in Arabic characters, in which at least certain basic letters which can have four different shapes correspond each to only one key of the keyboard; that when using said key one determines the shape to be used of the basic letters according to the characteristics of the keys used just after, or just before and after the said key, the preceding key used to select the shape of the desired letter being a key corresponding either to a letter than is linked with a following letter or to any other letter or sign or function which is not linked with a following letter, and the following key used to select the shape of the desired letter being a key which represents either any letter whether linked or not or any other sign or function which is not a letter. 35

40 2. The method claimed in claim 1 utilizing a keyboard as shown in Figure 2, in which a particular shape to be used for basic letters corresponding to keys 15, 26, 27, 30, 31 and 36 is determined by the characteristics of the two following keys when the first following key is one of the keys 26, 36 and by the characteristics of the first of said following key when the first key is 37, 39 or 44. 40

45 3. The method claimed in claim 1 or claim 2, in which the composition of the sentence BISM ILLAH ALROUHMAL AL RAHIM is automatically obtained by the use of only one particular key. 45

4. The method claimed in any preceding claim, when used for the transmission of a telex message using an existing network, in which one determines at reception if the signal received corresponds to a text in Arabic or not and that one switches accordingly the working mode of the receiver.

50 5. The method claimed in claim 4, in which this determination is made by storing the signals corresponding to at least one line of text and determining the presence or the absence in the stored signals of the sequence of codes corresponding to the function "changes of line" in the CCITT No. 2 Arabic alphabet which corresponds to a sequence of codes not used in the transmission of Latin characters. 50

55 6. The method claimed in claim 4, in which one prints automatically the service signals received from the exchanges with their translation in Arabic letters. 55

7. A method for the composition of texts in Arabic characters, substantially as hereinbefore described with reference to Figures 2 to 6 of the accompanying drawings.

60 8. Apparatus for the composition of texts in Arabic characters, comprising a keyboard having several keys each corresponding to only one basic character in all its four different shapes; an electronic device having a memory retaining the codes corresponding to all the shapes of each basic character and a working memory retaining the characteristics of the keys depressed immediately before and after the considered key; and a logic circuit for determining, according to the characteristics memorized in the working memory, the codes corresponding to the appropriate shape of the basic character, whereby said apparatus reproduces only that shape of a basic character that corresponds to the keys depressed 60



20

5

GB 2 013 577 B

5

immediately before and after the key corresponding to the basic character.

9. Apparatus as claimed in claim 8, in which the keyboard is electrically connected with an electronic circuit having a keyboard encoder, an input/output port for receiving signals from the encoder, a control device, a working flag register, a memory and an output device.

5

10. Apparatus as claimed in claim 9, in which the memory stores all the output codes corresponding to all the shapes or signs to be printed.

5

11. Apparatus as claimed in claim 10, in which the working flag register stores the characteristics of an output code already transmitted, or the characteristic of the character to be selected as well as its keyboard code; and the input/output port stores temporarily the keyboard of the last depressed key.

10

12. Apparatus for the composition of texts in Arabic characters, substantially as hereinbefore described with reference to and as illustrated in Figures 2 to 6 of the accompanying drawings.

10