(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
21 May 2015 (21.05.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/073766 A1

(51) **International Patent Classification:**
G10L 15/22 (2006.01) *G06F 17/27* (2006.01)

(21) **International Application Number:**
PCT/US2014/065597

(22) **International Filing Date:**
14 November 2014 (14.11.2014)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
1320334.4   18 November 2013 (18.11.2013)   GB
14/180,957   14 February 2014 (14.02.2014)   US

(71) **Applicant: MICROSOFT TECHNOLOGY LICENS-ING, LLC** [US/US]; One Microsoft Way, Redmond, WA 98052 (US).

(72) **Inventors: PLUMB, Graham**; c/o Microsoft Corporation, LCA - International Patents (8/1172), One Microsoft Way, Redmond, Washington 98052-6399 (US). **SWIFT, Mark**; c/o Microsoft Corporation, LCA - International Patents (8/1172), One Microsoft Way, Redmond, Washington 98052-6399 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:
— *with international search report (Art. 21(3))*

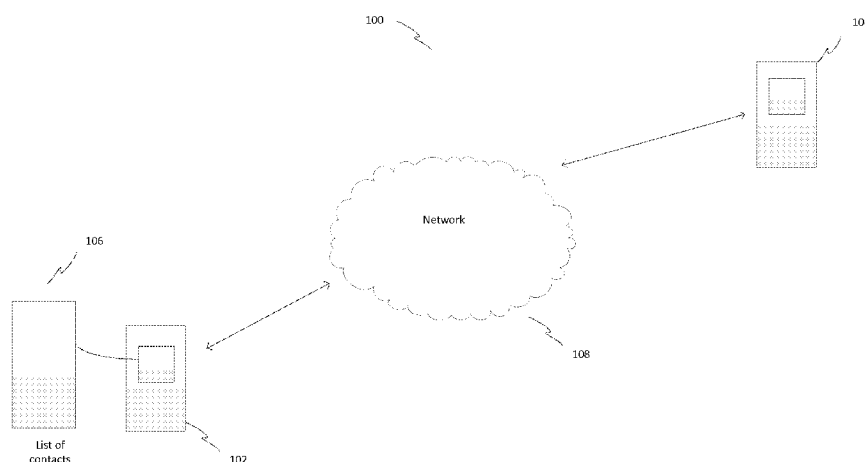(54) **Title:** IDENTIFYING A CONTACT

Figure 1

(57) **Abstract:** A method of identifying a contact in a communication system using voice input, the method comprising: receiving an input string of characters, the input string representing a contact and being normally unpronounceable by a human voice when spoken literally; performing at least one transforming step to transform at least one character of the input string to thereby generate a pronounceable name for the contact; and outputting the pronounceable name for use in establishing a communication event with the contact using voice input.

IDENTIFYING A CONTACT

BACKGROUND

[0001]   A contact's identification on services such as Skype and Xbox Live often consists of metadata like a username, gamertag, nickname, first name, surname, or combinations thereof. Some user communities actually encourage a degree of "digital anonymity" in contact names, for example the moniker "S0meG1rl" is intended to be pronounced as "Some Girl", but is unpronounceable if read literally. This presents a challenge for software attempting to use speech recognition to easily identify a contact without a high degree of (tangential) user interaction.

[0002]   Communication systems attempt to address this by doing one (or more) of the following: force users to enter additional pronounceable metadata; ignore difficult looking names or do nothing so that names entered as, for example, "Chr1s" are simply not useable in voice recognized commands.

SUMMARY

[0003]   The subject matter pertains to a method. The method is a method of identifying a contact in a communication system using voice input. The method comprises receiving an input string of characters, the input string representing a contact and being normally unpronounceable by a human voice if spoken literally. The method further comprises performing at least one transforming step to transform at least one character of the input string to generate a pronounceable name for the contact. The pronounceable name is output for use in establishing a communication event with the contact using voice input.

[0004]   The subject matter also pertains to a corresponding computer readable medium and a user device.

[0005]   This summary is provided to introduce a selection of concepts in a simplified form that are further described in the Detailed Description below. This summary is not intended to identify key features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006]   For a better understanding of the apresent disclosure and to show how the same may be put into effect, reference will now be made, by way of example, to the following drawings in which:

[0007]   Figure 1 shows a schematic illustration of a communication system;

[0008]   Figure 2 shows a schematic illustration of the operative components of a user device;

[0009]   Figure 3 shows a schematic illustration of the functional modules of a user device for identifying a contact using voice input;

[0010]   Figure 4 shows a flow chart of a method of identifying a contact using voice input; and

[0011]   Figure 5 shows a further schematic illustration of the functional modules of a user device for identifying a contact using voice input.

DETAILED DESCRIPTION

[0012]   In the present disclosure, transformation of name data and monikers to first name, surname and nicknames (where appropriate) allow for a more natural identification of a contact using voice input. By maintaining a score (quality parameter) for each step in a transformation, confidence in the resulting data may be inferred.

[0013]   A transformation process is a series of discrete sanitization operations which modify the input string source representing name data (be it full name, first name, moniker etc.) in an attempt to remove common patterns of complexity and ambiguous string components in order to generate a pronounceable name.

[0014]   The transformation process may implement transformation steps carried out on an input string representing name data/monikers for a contact to attempt to interpret "leet speak". "Leet speak" is the substitution of letters with numbers and symbols in a contact's name. For example, the aim would be to transform "S0meG1rl" to the pronounceable word "Some Girl".

[0015]   Each transformation step returns a result that contains the source name data, transformed data and the number of character changes or removals undertaken. The output of each transformation is then chained to the input of the next transformation until all steps are complete and a final result is ascertained.

[0016]   The result can then be used to calculate a score (quality parameter) based on the size of the original text, and the cumulative transformations that have been undertaken upon it, allowing for evaluation against a quality threshold. The approach will not entirely mitigate the issue of incorrect pronunciation of complex monikers (such as, for example, Xbox Live gamertags), but will go some way to improving its likelihood.

[0017]   In one embodiment, the output of these steps is intended to be used as part of a speech recognition system's grammar, so there will not be a visual component (the

user will experience on the display the original source text). However, in other embodiments, the pronounceable name is outputted to the display and displayed in association with the input string. In some embodiments, the pronounceable name may be outputted to the display in editable form as an autosuggested replacement for the input

5   string.

[0018]   Optionally, the sanitization operations of the transformation process are based on a contact's region. This may allow for cultural differences in, for example, "leet speak", contact titles and/or honorifics. It should be noted that each step of the transformation process may perform regional sanitization.

10   [0019]   Figure 1 shows a general communication system 100 comprising a first user terminal 102 associated with a first user (User A, not shown) and a second user terminal 104 associated with a second user (User B, not shown). Whilst only two user terminals have been shown in Figure 1 for simplicity, the communication system 100 may comprise any number of users and associated user devices. The user terminals 102 and

15   104 can communicate over the network 108 in the communication system 100, thereby allowing the users A and B to communicate with each other over the network 108. The network 108 may be any suitable network which has the ability to provide a communication channel between the first user terminal 102 and the second user terminal 104. For example, the network 106 may be the Internet or another type of network such as

20   a High data rate mobile network, such as a 3rd generation ("3G") mobile network. The user terminal 102 may be, for example, a mobile phone, a personal digital assistant ("PDA"), a personal computer ("PC") (including, for example, Windows™, Mac OS™ and Linux™ PCs), a gaming device or other embedded device able to connect to the network 108. The user terminal 102 is arranged to receive information from and output information to the

25   user of the user terminal 102. In one embodiment the user terminal 102 comprises a display such as a screen and an input device such as a keypad, a touch-screen, and/or a microphone. The user terminal 102 is connected to the network 108.

[0020]   Each user terminal 102 and 104 may use a contact list to store the contact details of other user terminals associated with other users. For example, in the general

30   communication system shown in figure 1, user terminal 102 associated with User A uses a contact list 106. The contact list 106 could store contact information of other users with which User A may wish to establish a communication event, for example the contact list could store the contact details of user terminal 104 associated with User B. The contact information could be, for example, a telephone number, an email address, an IP address

etc. The contact list 106 associates a contact's identification with their contact information. The contact's identification could be, for example, the name of the contact, a username, their gamertag, a nickname or combinations thereof.

[0021]   The contact list is accessible by User A in order to enable User A to establish a communication event with a user stored on the list. For example, the contact list may be capable of being displayed on a graphical user interface of the user terminal. The user could then select a relevant contact from the displayed contact list in order to establish a communication event with the terminal of the selected contact. The contact list could also be accessed via a voice input from the user. For example, the user could pronounce the name of a contact on the contact list and the user terminal could process this voice input to select the relevant contact from the contact list and establish a communication event with the associated user terminal.

[0022]   Figure 2 illustrates a view of the operative components of a user terminal 102. It should be noted that as used herein like reference numerals represent like components/elements. The user terminal comprises a microprocessor 204 connected to a display 205 such as a screen or touch screen. The processor is further connected to an audio input device 206 (e.g. a microphone) and an audio output device 207 (e.g. a speaker). The display 205, audio input device and audio output device may be integrated into the user terminal 102 as shown in figure 2. In alternative user terminals one or more of the display, audio input and audio output may not be integrated into the user terminal 102 and may connect to the microprocessor via respective interfaces. The microprocessor is connected to a network interface 208 such as a modem for communication with the network 105. The network interface 208 may be integrated into the user terminal 102 as shown in Figure 2. In alternative user terminals the network interface 208 is not integrated into the user terminal 102. The microprocessor is further connected to an Automatic Speech Recognition (ASR) engine 209. The user terminal 102 comprises a memory 201 having stored thereon a Grammar 202, transform logic 203 and the contact list 106. The Grammar is connected to the ASR engine 209. The memory 201 may be a permanent memory, such as ROM, or may alternatively be a temporary memory, such as RAM.

[0023]   The ASR engine 209 is coupled to the microphone 206 and is configured to receive from the microphone an input audio signal, such as a voice signal generated from a user speaking into the microphone. The ASR engine is configured to analyse the input audio signal to determine if speech can be recognised and whether that speech represents an instruction from the user.

[0024]   The ASR engine determines whether speech can be recognised by accessing the Grammar 202. For example, the user may speak into the microphone to instruct the user terminal to call a contact from the contact list 106, such as User B. The microphone will generate from the user's instruction an audio signal. The ASR engine will analyse this signal by accessing the Grammar, and if it recognizes the name of the contact spoken by the user (in this case User B), it will output a signal to trigger the user terminal 102 to call terminal 104 associated with User B.

[0025]   In some instances the contact list contains contacts whose identification is not pronounceable when read literally by a human. Such an example is the "leet speak" described above. Contact names written in "leet speak" can be written in such a way that, although not pronounceable when read literally, can nevertheless be understood when read by a human. Referring to the earlier example above, the moniker "S0omeG1rl" can be understood as "Some Girl" when read by a human even though the moniker itself it not pronounceable.

[0026]   Maintaining this example, if User A speaks into the microphone of terminal 102 to request the terminal to call "Some Girl" (intending for terminal 102 to call the terminal associated with the contact "S0meG1rl"), the ASR engine will recognise this speech, but will not be able to trigger the user terminal 102 to call the desired terminal because there is no contact identifier on the contact list 106 under the name "Some Girl". In these instances the ASR engine may not be successful in using speech recognition to identify a contact.

[0027]   Referring to figures 3 and 4 there is described a method for transforming an input string into a human-pronounceable word or name. The input string could be, e.g., name data or a moniker, or be written in "leet speak". The input string could represent a contact's identification for use within a contact list.

[0028]   Figure 3 shows an input string 301 which is received at an input of the transform logic 203 and a store module 302. The transform logic has an output configured to output transformed data 303, which could be a word pronounceable by a human voice. The transform logic 203 is configured to perform the transformation process from the input string to the transformed data. As will be described in more detail below, the transformation process could comprise multiple transformation steps which modify the input string in an attempt to remove common patterns of complexity and ambiguous string components. Each of these steps could be implemented by the transform logic 203.

[0029]   In one embodiment the transform logic is a software program executed on a local processor in the user terminal, such as processor 204. The transform logic may have an input configured to receive an input string, and an output configured to output transformed data.

[0030]   Further output from the transform logic is Quality Indicator (QI) 304 which is a score used to evaluate the quality of the transformation. The score can be calculated based on the size of the original input string and the transformations that have been undertaken on it to produce the transformed data.

[0031]   If the transformation process comprises multiple transformations steps, the QI can be calculated as a function of a calculated result for each of the transformation steps. For example, each transformation step could return a result that contains the input string (for that step) and information relating to the degree of transformation between the input string and the transformed data for that step, such as the number of character changes undertaken to transform the input string into the transformed data, and/or the number of characters removed from the input string. The output of each transformation step is then chained to the input of the next transformation step until all the transformation steps are complete and a final result is calculated. The QI is then calculated from the final result.

[0032]   A scoring module 305 is configured to receive the output transformed data and the QI. Based on the value of the QI, the scoring module can determine whether to discard the transformed data 303 or to output the transformed data for addition to the Grammar 202. For example, the scoring module could compare the calculated score QI to a quality threshold. If the score is below the threshold, the transformed data is discarded; if the score is above the threshold, the transformed data is added to the Grammar 202.

[0033]   In one embodiment the scoring module is a software program executed on a local processor of the device, such as processor 204.

[0034]   Further shown in figure 3 is the ASR engine 209 which is configured to be able to access the Grammar 202. The ASR engine is configured to receive as an input an audio signal 306 (such as an audio signal output from the microphone 206) and to output a command function 307 for triggering the user device to perform an operation based on the input audio signal. The audio signal could be a command to, for example, call, email or message a contact from the contact list 106.

[0035]   In one embodiment, the source input strings are a contact's identification from a user's contact list 106 stored on a user terminal 102. In this embodiment the output

transformed data 303 could be a human-pronounceable name or word that is associated with the contact's identification. Referring to our earlier example, the contact's identification could be the moniker "S0meG1rl". In this example the source input string is data representing this moniker and the output transformed data could be data representing

5   the pronounceable name "Some Girl". If the transformed data is of sufficient quality the name "Some Girl" would be added to the Grammar 202 to be associated with the moniker "S0meG1rl".

[0036]   With reference to figure 4 there is described a method of identifying a contact from the contact list using a voice input.

10   [0037]   At step 401 a source input string 301 is received at transform logic 203. The input string is a contact's identification from the contact list 106. The transform logic could be configured to receive the input string from the memory 201 of the user device which stores the contact list 106. The contact's identification could be written in leet speak or be a moniker, and as such be unpronounceable by a human voice if spoken

15   literally.

[0038]   At step 402 at least one transformation step is performed on the input string to transform at least one character of the input string. The at least one transformation step is performed by the transform logic 203. During each transformation step, characters of the input string are transformed, e.g. by substitution or by deletion, in order to remove

20   characters of the input string that contributed to the string being unpronounceable by a human voice.

[0039]   After completion of the at least one transformation step, the unpronounceable input string has been transformed into a pronounceable name for the contact's identification. The pronounceable name is output from the transform logic 203

25   in step 403 as transformed data.

[0040]   In one embodiment, the transformed data is input into a scoring module 305 to evaluate the quality of the transformed data, as described above. In this embodiment, if the scoring module determines that the transformed data is of sufficient quality (e.g. compared to a threshold quality), the transformed data is input into the

30   Grammar 202 of the user terminal. If the transformed data is not of sufficient quality, it is discarded and not input into the Grammar.

[0041]   In an alternative embodiment, the transformed data is output from the transform logic into the Grammar without evaluating its quality. That is, all transformed data output from the transform logic will be input into the Grammar.

[0042]   By inputting the transformed data into the Grammar 202, the contact's identification is then associated with a pronounceable name. If the user terminal receives a voice input from the user which contains the pronounceable version of a contact's identification, the ASR engine will be able to access the Grammar 202 to recognise the pronounceable name and determine that it is associated with a contact's identification from the contact list. Based on the command input by the user, the ASR engine can then output a message to trigger the user terminal to take an appropriate action. For example if the voice input command was "Call 'Some Girl'", the ASR engine would access the Grammar to recognise the name "Some Girl" and identify that this name is associated with the identification "S0meG1rl" from the contact list 106. The message output from the ASR engine would then trigger the user terminal to call the contact "S0meG1rl".

[0043]   In one embodiment, the source input string identifying the contact from the contact list is output to the display 205 of the user terminal. For example, a message may be output to the display to inform the user that the displayed source input string is going to be transformed into a pronounceable name. Alternatively a message may be output to ask the user whether they wish for the displayed source input string to be transformed into a pronounceable name, and to provide the user with the option of not proceeding with the transformation if they do not wish to do so.

[0044]   Alternatively or in addition, a message may be output to the display of the user terminal once the at least one transformation step has been performed on the source input string to display the transformed data to the user. This message could contain both the source input string and the associated transformed data. For example, a message could be output to inform the user that the source input string has been transformed into a human-pronounceable word. This could make the user aware that the pronounceable word has been added to the Grammar and will be recognised by the ASR engine. In another embodiment, the transformed data could be output to the display of the user terminal in the form of an editable autosuggestion. This could allow the user to edit the transformed data if the input string has been transformed incorrectly, for example because the transformed data is not pronounceable.

[0045]   As described above, the transformation performed by the transform logic 203 to transform the source input string to the transformed output data could comprise more than one transformation step. In one embodiment, the operations performed by the transform logic in a transformation are as follows:

[0046]   *Replace accents:* This step changes accented characters to their base equivalents where regionally acceptable (for example, the transformation of á to a is an acceptable transformation in UK English but not in French).

[0047]   *Strip symbols:* Remove inappropriate symbols such as, for example, © and ™.

[0048]   *Change separators to whitespace:* This step removes characters in the input string such as dashes, underscores, hyphens and other adjoining symbols.

[0049]   *Trim leading/trailing numbers and symbols:* This step removes any leading or trailing numbers from the input string (e.g. Mark23 to Mark).

[0050]   *Replace double vowel leet characters:* This step transforms double-vowel leet characters into their appropriate equivalents, for example 00 to oo, 33 to ee etc.

[0051]   *Replace single vowel leet characters:* This step changes leet characters that exist within a run of characters, optionally evaluating preceding characters. For example, this step would transform: S1mon to Simon, A1an to Alan, Mich@el to Michael etc.

[0052]   *Strip initials:* This step removes single characters that end with a period or whitespace, for example Joe M Bloggs transforms to Joe Bloggs.

[0053]   *Normalize capitalisation:* This step detects runs of characters where name fragments are combined and only distinguishable by capitilisation. This may occur in, for example, gamertags. Once detected, this step inserts whitespace to separate the name fragments, e.g. JohnPaulSmith is transformed to John Paul Smith.

[0054]   *Parse suffixes, titles, nicknames and fullname:* This step parses any suffixes or titles (e.g. Mr, Mrs, PhD etc.) and classifies them against a known list. The left-over characters are then parsed into a first name (the first token in the string, that is, the first series of contiguous characters without a line break) and a full name (the entire string, once whitespace has been compacted).

[0055]   Figure 5 is a schematic diagram to illustrate the modules of a user device for identifying a contact in a communication system using a voice input.

[0056]   Figure 5 shows a contact list 106 which stores details of contacts of an end user 502. The contact list could be stored on a user terminal, such as terminal 102, which is operable by end user 502. Each contact on the contact list will have an associated contact name and contact information (e.g. a phone number, email address, IP address or a combination thereof). The contact name of contact 501 could be unpronounceable by a human voice when spoken literally. An example of such a contact name is the moniker "Chr1s", which will be used herein in relation to figure 5 for the purposes of illustration

only. The contact list could be accessible by the user 502, for example by display on a graphical user interface of the user terminal.

[0057]   The contact name of contact 501 is input in the form of a source input string 301 into a transform logic 203. The transform logic is configured to perform at least one transformation step on the input string. The transformation steps performed by the transform logic could include any and all of the operations described above. The output of the transform logic is transformed data 303. The transformed data 303 is a transformed version of the contact name that could be pronounceable by a human when spoken literally. For example, for the contact name "Chr1s", the transformed data could be the name "Chris".

[0058]   The transformed data is input into a Grammar 202 so that the pronounceable name "Chris" is associated with the contact name "Chr1s".

[0059]   In one embodiment, the contact names from contact list 106 could be input into the transform logic module automatically, that is to say, without user input. For example, the user may enter the details of a contact into the contact list 106. Upon entering the details, the contact name is input into the transform logic module. In an alternative embodiment, the contact names could be input into the transform logic module to be transformed into pronounceable names upon the request of the user. In yet another alternative embodiment, the user terminal could request permission from the user to transform the contact names of contact list 106 into pronounceable names. If the user refuses permission, names from the contact list are not input into the transform logic module 203.

[0060]   Once a contact name has been transformed by logic module 203 and stored in the Grammar 202, the contact can be identified by the user 502 from a voice input 503. For example, the user may speak a command containing the word "Chris". An automatic speech recognition (ASR) engine 309 is configured to receive the voice input and to access the grammar 202. By accessing the grammar, the ASR engine can recognise the word "Chris" and determine that it is associated with the contact name "Chr1s". The ASR engine then outputs a command 504 to trigger an action associated with the contact name "Chr1s", such as triggering the user terminal to establish a communication event with the user's contact "Chr1s".

[0061]   It will be understood that a transformation process may contain any number of the above steps. For a process containing more than one step, the input string is passed through each step sequentially, such that the output of one step is supplied as the input to

the next step. In one embodiment, the transformation process comprises all of the operations described above. The steps may be performed in the order in which they are described, e.g. the first step is to replace accents, the second step (performed on the output of the first step) is to strip symbols, and so on.

[0062]   Generally, any of the functions described herein can be implemented using software, firmware, hardware (e.g., fixed logic circuitry), or a combination of these implementations. The terms "module," "functionality," "component" and "logic" as used herein generally represent software, firmware, hardware, or a combination thereof. In the case of a software implementation, the module, functionality, or logic represents program code that performs specified tasks when executed on a processor (e.g. microprocessors, CPU or CPUs). The program code can be stored in one or more computer readable memory devices. The features of the techniques described below are platform-independent, meaning that the techniques may be implemented on a variety of commercial computing platforms having a variety of processors.

[0063]   For example, the user devices may also include an entity (e.g. software) that causes hardware of the user devices to perform operations, e.g., processors functional blocks, and so on. For example, the user devices may include a computer-readable medium that may be configured to maintain instructions that cause the user devices to perform operations. Thus, the instructions function to configure the operating system and associated hardware to perform the operations and in this way result in transformation of the operating system and associated hardware to perform functions. The instructions may be provided by the computer-readable medium to the user devices through a variety of different configurations.

[0064]   One such configuration of a computer-readable medium is signal bearing medium and thus is configured to transmit the instructions (e.g. as a carrier wave) to the computing device, such as via a network. The computer-readable medium may also be configured as a computer-readable storage medium and thus is not a signal bearing medium. Examples of a computer-readable storage medium include a random-access memory (RAM), read-only memory (ROM), an optical disc, flash memory, hard disk memory, and other memory devices that may use magnetic, optical, and other techniques to store instructions and other data.

[0065]   Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts

described above. Rather, the specific features and acts described above are disclosed as example forms of implementing the claims.

CLAIMS

1.   A method of identifying a contact in a communication system using voice input, the method comprising:

receiving an input string of characters, the input string representing a contact and being normally unpronounceable by a human voice when spoken literally;

performing at least one transforming step to transform at least one character of the input string to thereby generate a pronounceable name for the contact; and

outputting the pronounceable name for use in establishing a communication event with the contact using voice input.

2.   A method as claimed in claim 1, wherein the input string is displayed to a user on a display of a user terminal, wherein the pronounceable name is outputted to the display in editable form as an autosuggested replacement for the input string.

3.   A method as claimed in claim 1, wherein the pronounceable name is stored in a grammar accessible by a speech recognition engine, whereby voice input of a pronounced version of the input string to the speech recognition engine locates a match of the pronounceable name.

4.   A method as claimed in claim 1 further comprising the steps of:

outputting a quality parameter indicating the quality of the relationship between the input string and the pronounceable name; and

discarding pronounceable names below a quality threshold, using the quality parameter.

5.   A method as claimed in claim 1 wherein the at least one transforming step comprises multiple transforming steps, the output of one step being supplied to a next step wherein the at least one transforming step comprises at least one of: replacing accents, stripping symbols, changing separators to whitespace, trimming leading/trailing numbers and symbols, replacing double vowel leet characters, replacing single vowel leet characters, stripping initials, normalising capitalisation, inserting whitespace in contiguous runs of characters between fragments distinguished by capitalisation, parsing suffixes and/or titles and/or nicknames and/or fullnames, and classifying said parsed suffixes and/or titles and parsing remaining characters of the input string into a first name and a fullname.

6.   A method as claimed in claim 5, wherein the at least one transforming step comprises all of the transforming steps of claim 5.

7.   A method as claimed in claim 1, wherein the input string representing the contact is written in "leet speak".

8.   A method as claimed in claim 4, wherein the at least one transforming step comprises multiple transforming steps, the output of one step being supplied to a next step and wherein the quality parameter is dependent upon a score for each step of the said multiple transforming steps.

9.   A computer readable medium storing code for identifying a contact in a communication system using voice input which, when executed on a processor of a user device, is configured to:

        receive an input string of characters, the input string representing a contact and being normally unpronounceable by a human voice when spoken literally;

        perform at least one transforming step to transform at least one character of the input string to thereby generate a pronounceable name for the contact; and

        output the pronounceable name for use in establishing a communication event with the contact using voice input.

10.   A user terminal configured to identify a contact in a communication system using voice input, the user terminal comprising:

        an input configured to receive an input string of characters, the input string representing a contact and being normally unpronounceable by a human voice when spoken literally;

        transform logic configured to perform at least one transforming step to transform at least one character of the input string to thereby generate a pronounceable name for the contact;

        an output configured to output the pronounceable name for use in establishing a communication event with the contact using voice input and output a quality parameter indicating the quality of the relationship between the input string and the pronounceable name; and

        a scoring module configured to discard pronounceable names below a quality threshold, using the quality parameter.
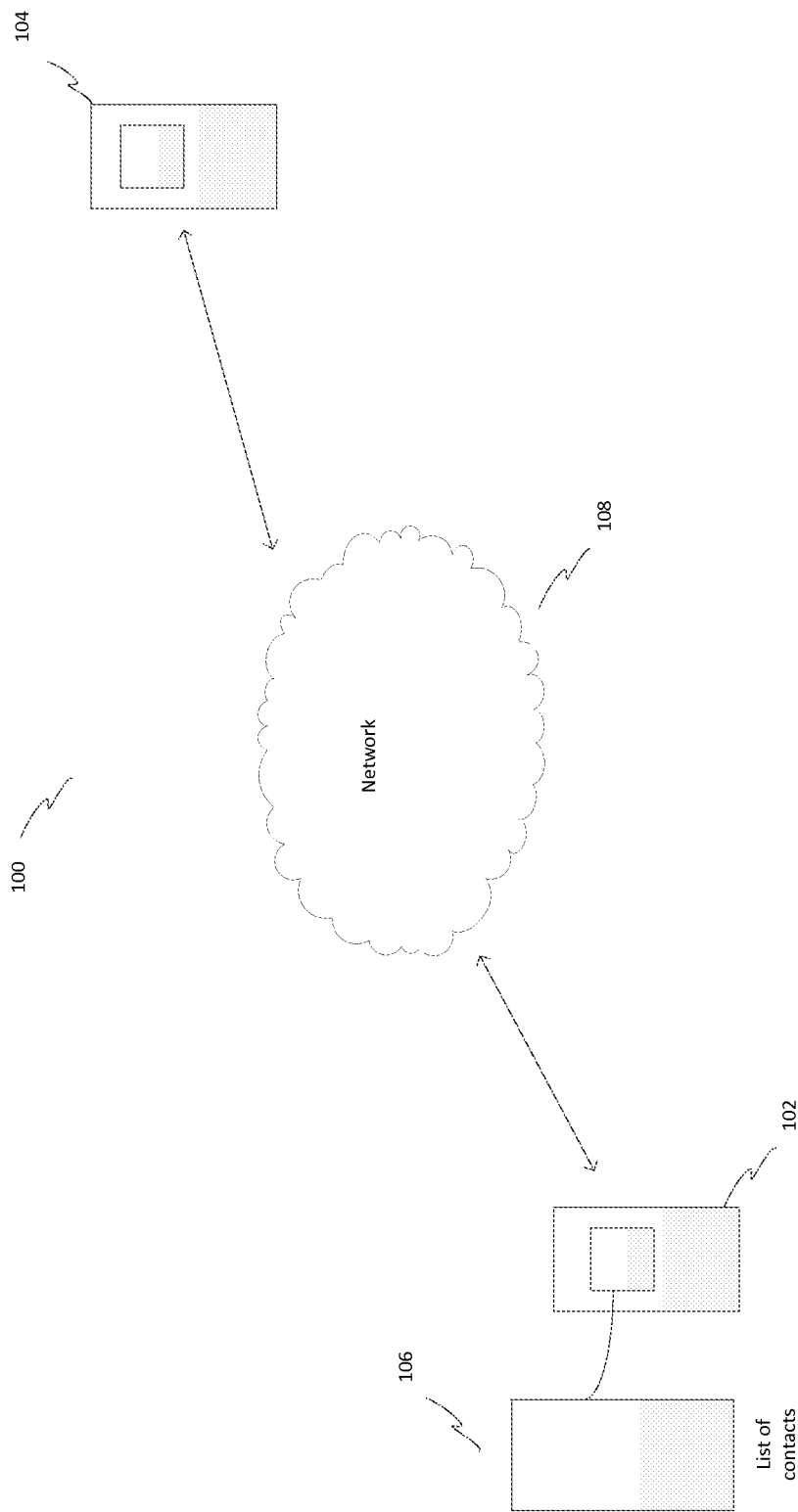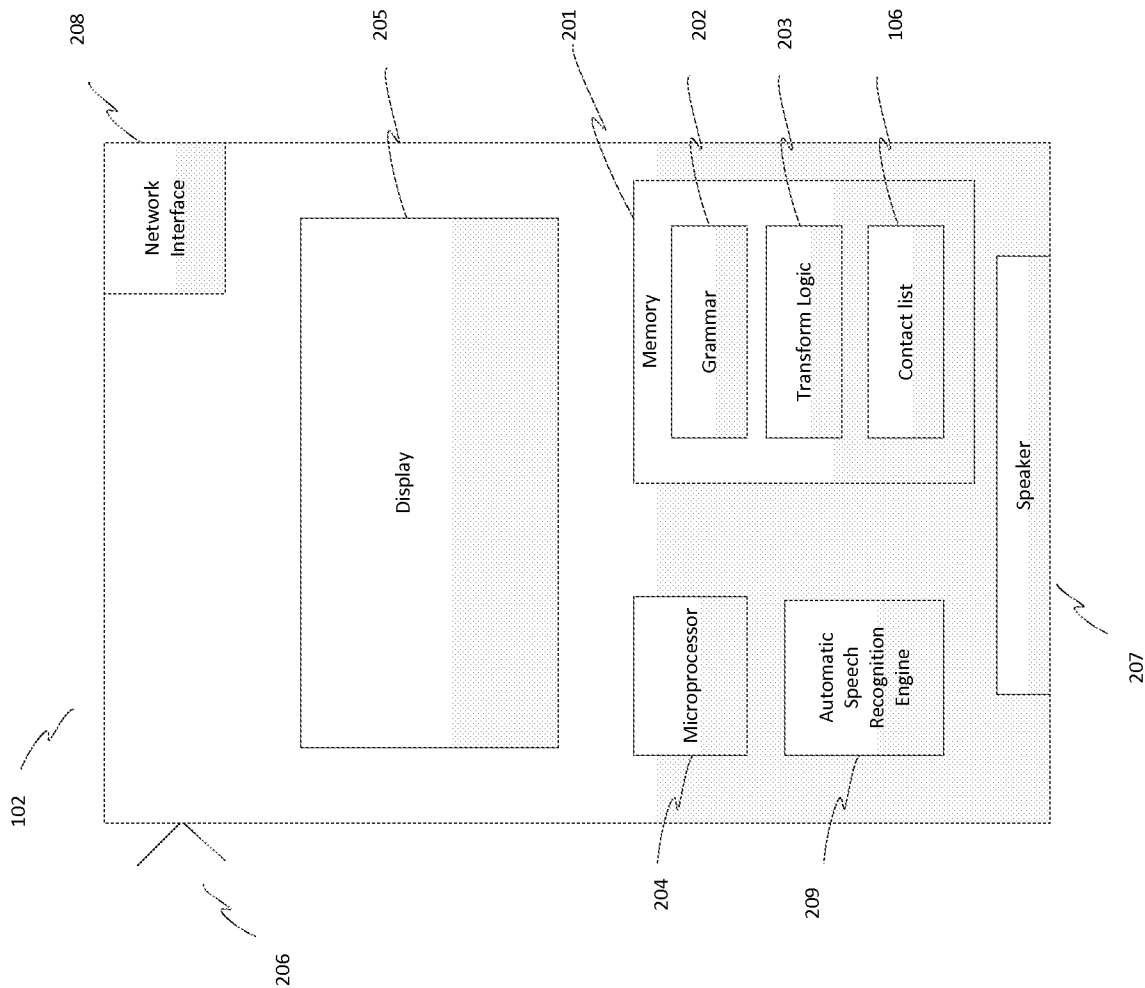
Figure 1

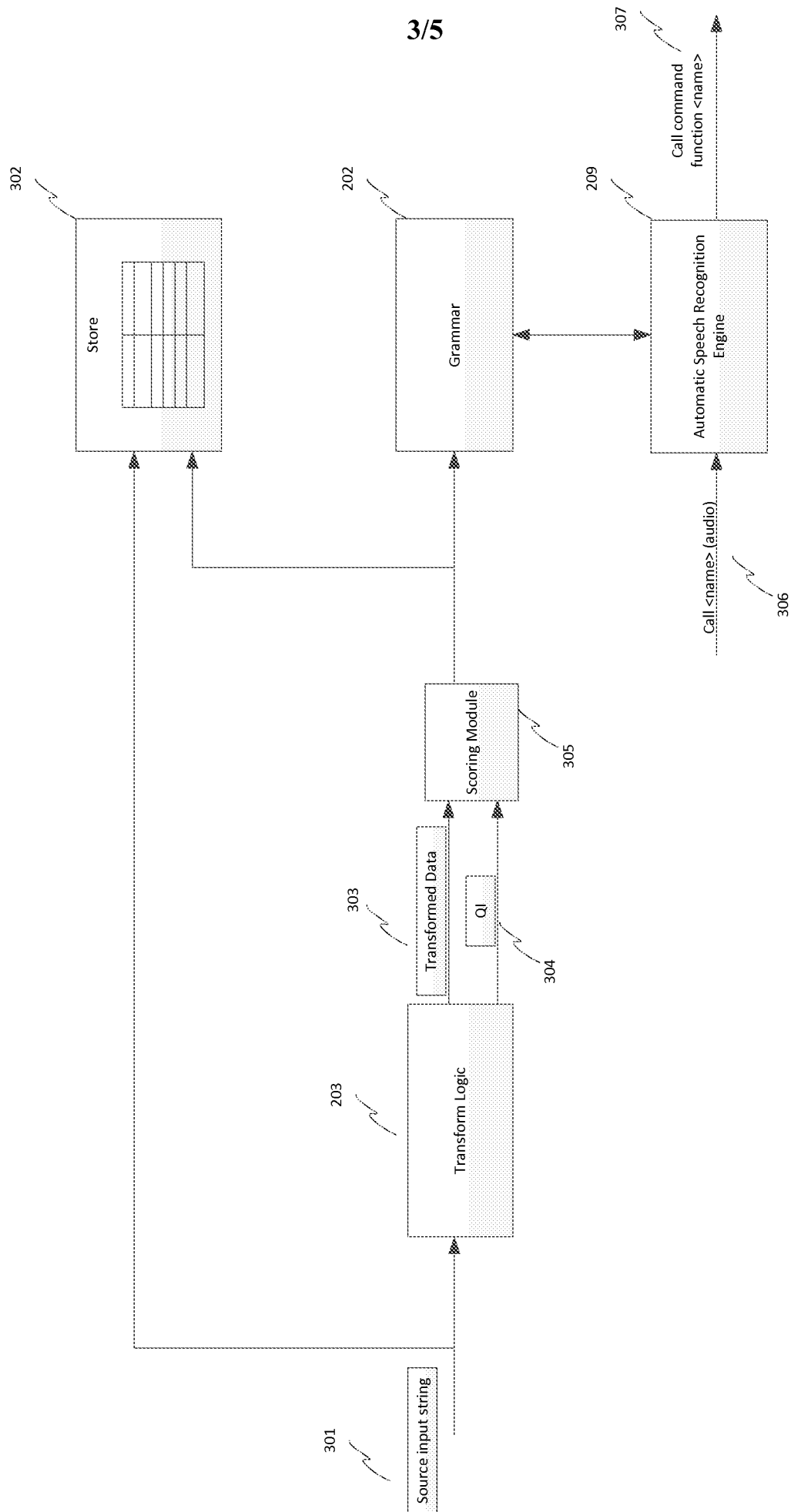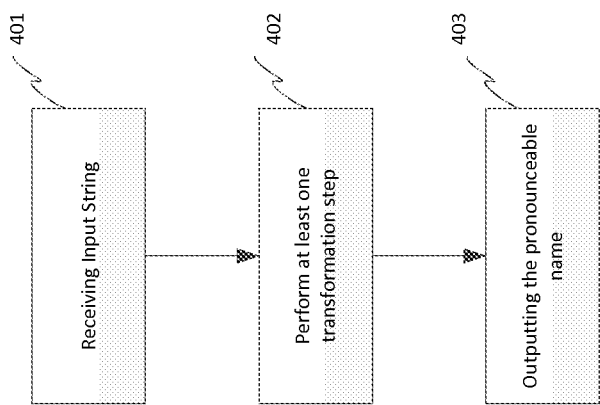Figure 2

Figure 3

401

Receiving Input String

402

Perform at least one
transformation step

403

Outputting the pronounceable
name
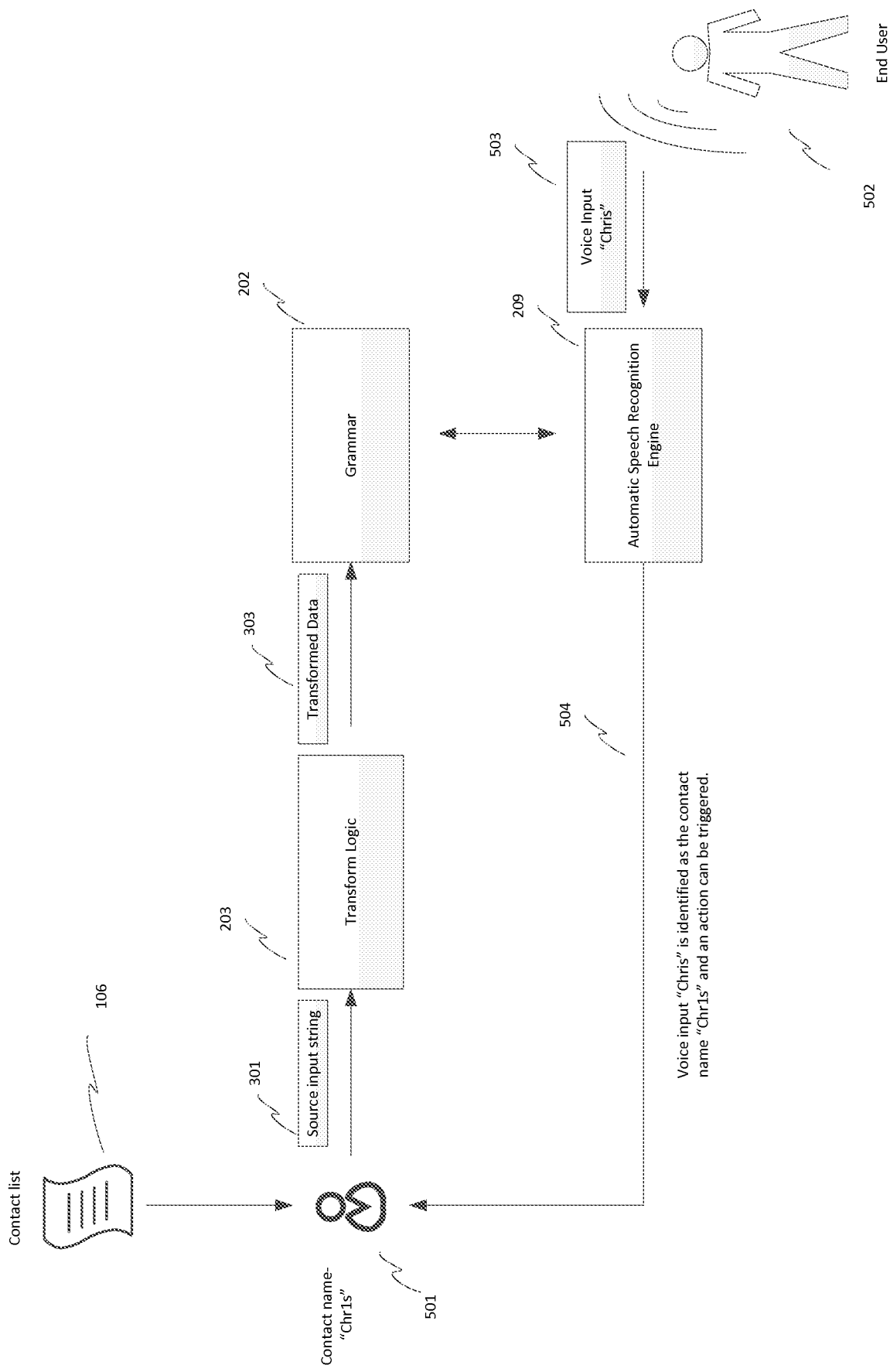
Figure 4

Figure 5

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
INV. G10L15/22   G06F17/27
ADD.

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)
G10L G06F

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)

EPO-Internal, COMPENDEX, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013/231917 A1 (NAIK DEVANG K [US]) 5 September 2013 (2013-09-05) the whole document -----	1-10
X	US 7 467 087 B1 (GILLICK LAURENCE S [US] ET AL) 16 December 2008 (2008-12-16) claims 1,5,8 column 2, line 46 - line 61 -----	1-10
X	US 6 314 165 B1 (JUNQUA JEAN-CLAUDE [US] ET AL) 6 November 2001 (2001-11-06) the whole document ----- -/--	1-10

[X] Further documents are listed in the continuation of Box C.　　　　[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 February 2015	04/03/2015

| Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016 | Authorized officer Deltorn, Jean-Marc |

2

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	XIAO LI ET AL: "Adapting grapheme-to-phoneme conversion for name recognition", AUTOMATIC SPEECH RECOGNITION&UNDERSTANDING, 2007. ASRU. IEEE WORKS HOP ON, IEEE, PI, 1 December 2007 (2007-12-01), pages 130-135, XP031202047, DOI: 10.1109/ASRU.2007.4430097 ISBN: 978-1-4244-1745-2 sections 1 and 2	1-10

X	BEAUFAYS F ET AL: "Learning name pronunciations in automatic speech recognition systems", PROCEEDINGS 15TH IEEE INTERNATIONAL CONFERENCE ON TOOLS WITH ARTIFICIAL INTELLIGENCE. ICTAI 2003. SACRAMENTO, CA, NOV. 3 - 5, 2003; [IEEE INTERNATIONAL CONFERENCE ON TOOLS WITH ARTIFICIAL INTELLIGENCE], LOS ALAMITOS, CA, IEEE COMP. SOC, US, vol. CONF. 15, 3 November 2003 (2003-11-03), pages 233-240, XP010672233, DOI: 10.1109/TAI.2003.1250196 ISBN: 978-0-7695-2038-4 the whole document	1-10

A	Zhenzhen Xue ET AL: "Normalizing Microtext", Analyzing Microtext 11: 05., 1 January 2011 (2011-01-01), XP055171754, Retrieved from the Internet: URL:http://www.researchgate.net/profile/Brian_Davison/publication/221603856_Normalizing_Microtext/links/00b4951e732d911d99000000.pdf [retrieved on 2015-02-24] the whole document	1-10

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 2013231917	A1	05-09-2013	AU 2013225826	A1	18-09-2014
			EP 2815397	A2	24-12-2014
			KR 20140136969	A	01-12-2014
			TW 201403585	A	16-01-2014
			US 2013231917	A1	05-09-2013
			WO 2013130878	A2	06-09-2013
US 7467087	B1	16-12-2008	NONE		
US 6314165	B1	06-11-2001	CN 1246023	A	01-03-2000
			DE 69922872	D1	03-02-2005
			DE 69922872	T2	29-12-2005
			EP 0953967	A2	03-11-1999
			JP 3135233	B2	13-02-2001
			JP 2000032140	A	28-01-2000
			TW 424221	B	01-03-2001
			US 6314165	B1	06-11-2001