Georgian Speech Recognizer in Famous Searching Systems and Management of the Software Package by Voice Commands in Georgian Language

Irakli Kardava

Abstract— **The speech recognition system and creation of program Software based on it is one of the important components of natural language computer modelling. For example, the popular Google Voice system, by which the word or sentence pronounced by a person is automatically written in textual form, in Google's Search system field for given language by using already realized speech recognition systems. The trial Software program has been developed by us, which is able to solve the problems, as for the Google search system, also for YouTube, Yahoo, Bing, Yandex and Wikipedia in Georgian language. Also, it is possible to translate automatically Georgian words or sentences pronounced individually and get texts in the appropriate language by using Google Translate. Program management is performed with Georgian voice commands.**

Keywords— **Georgian voice commands, Google voice, Software, Speech recognition.**

I.   Introduction

In general, the Software represents a web browser, connected to the transcription file [1] [2] created by us, base of words [3] (For composition of words we use the software, developed by us, it represents a set of programs, by which it is possible with unchanged part of a word and morphological categories to get appropriate grammatically right word-form or word-forms if such exist. Also with this software it is possible by unchanged part of a word get all possible grammatically right word-forms. This used approach is based on description of natural language morphology by using formal grammar and characterizing symbols of grammar with feature structures. For description natural language morphology we use special type of context free grammar, which describes all correct natural language word-forms. With given unchanged part and its features existed in database, also with given morphological categories (in the case of first problem) we compose morpheme classes and their representatives, which must be in related word-forms. Using of the software is effective for languages, which have developed morphology like Georgian.) [4] and acoustic model of the Georgian language. The constant cycle is engaged in browser, which dynamically checks the current processes in speech recognition system [5].

However, since the speech recognition system has been created, it has developed significantly, but it still has a lot of problems. As you know, any specific natural language may owns about tens accents.

Irakli Kardava
Sokhumi State University in Tbilisi, Georgia
Adam Mickiewicz University in Poznan, Poland

Despite the identical word phonemic composition, if it is pronounced in different accents, as a result, we will have sound waves, which are different from each other. Differences in pronunciation, in accent and intonation of speech in general, create one of the most common problems of speech recognition. For solving the problem of the accens, we have carried out an observation of an individual sound of vowels, consonants and studied each of these structures [6]. As a result, we have divided their sounds into I and II phases, which clearly showed that the second phases of consonants are equal to each other i.e. they are identical, and the difference between them is only the first phase (including the relevant Allophones) [7]. The changes in accents mainly are caused due to the elongation of the second phase in time, while the first phase during speaking remains unchanged. Therefore, it is enough to observe the sound of the first phase of speech recognition [8], because the extension of time of the second phase cannot change the phonemic composition of word or sentence. Similarly, it is achievable to resolve the problem [9] of the tempo of pronounced words or sentences [10] during speech recognition [11] [12]. (The paper about Solving the Problem of the Accents for Speech Recognition Systems is accepted and will be published soon in International Journal of Signal Processing Systems (IJSPS), USA.)

When program is started and is ready to use, in case of recognize some words and delivering the new text from speech recognizer, the information is sent to navigation function of the browser, with discovered text. Foreseen the fact, which search system is chosen by a user, the search or translation of designated phrase, has been implemented by using the same site. In case, if the phrase coincides with the words pre-reserved by us, the function for performance relevant statements will be activated, which finally gives us possibility to manage the Software with voice command, without using buttons of keyboard and mouse.

II.   The Software Design

The software represents a set of programs. One part of them works for speech recognition and second part works directly for using base possibilities of the software. The main window of above mentioned software, which is connected to the georgian speech recognizer also created by us, looks as follows, see the Figure 1.

In the upper part of the program interface, the names of some systems are enumerated. The users can use each of them while working the program. To activate the system, it is necessary to select the one from the list (By checking a given radiobutton), depending on the user's wish.
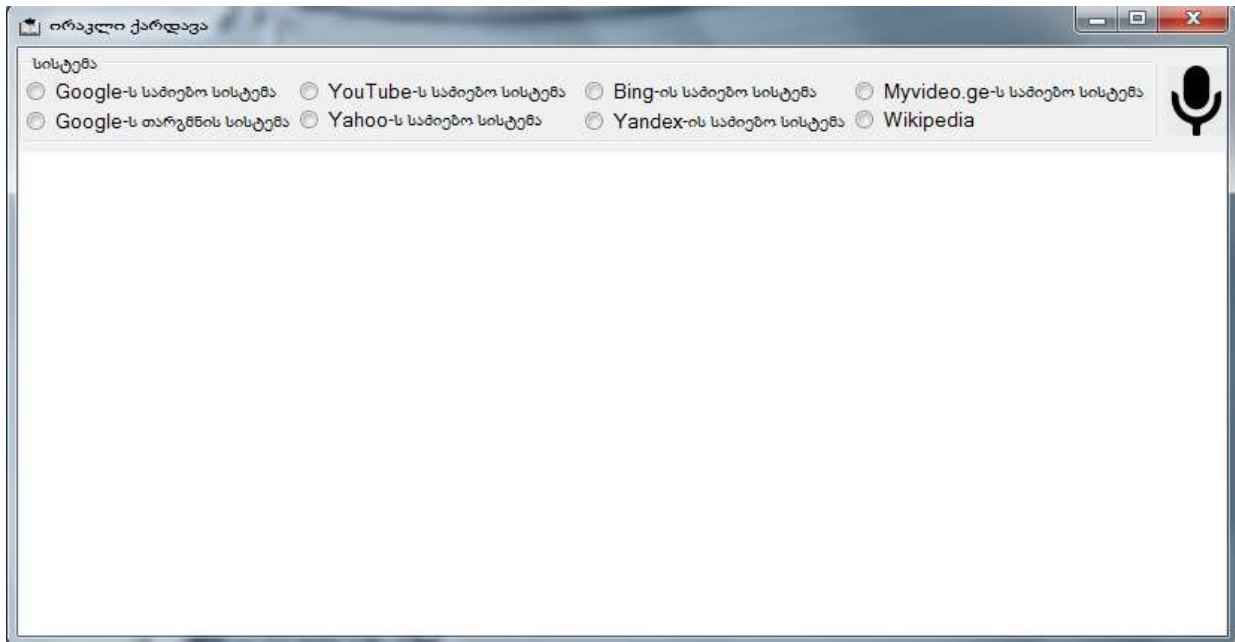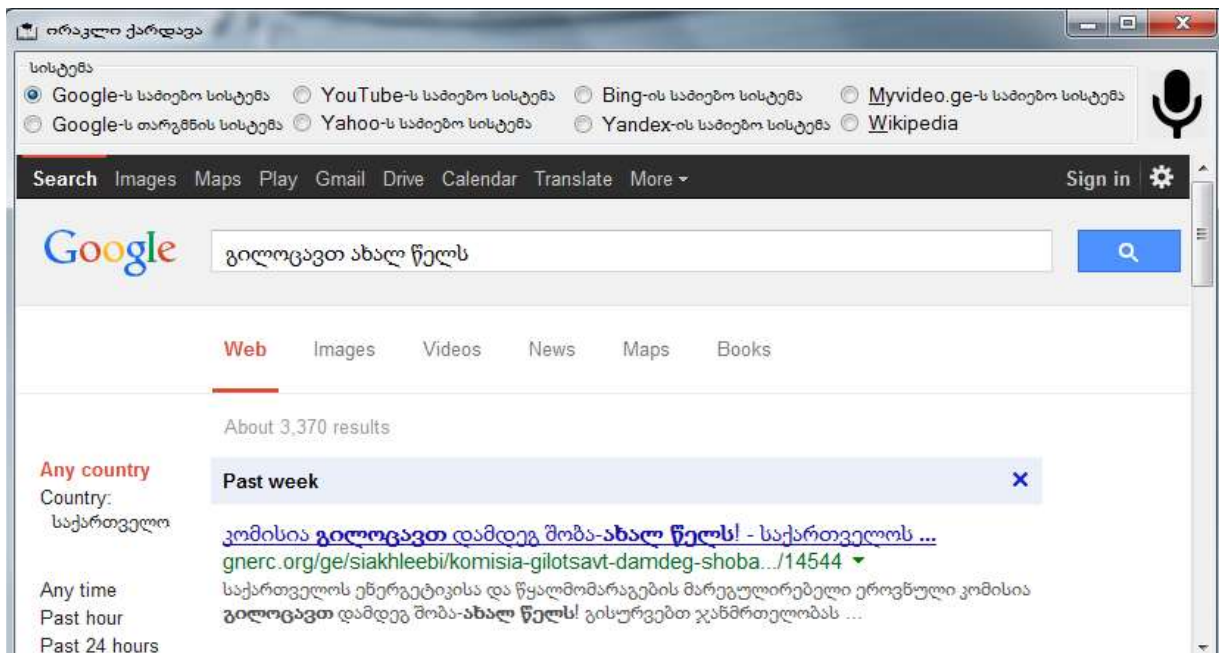
Figure 1. The main window of software.

Figure 2. The result of search according to the pronounced sentence.

For example, choose the Google search field, choose the category from the list "Google's search system" and pronounce the following sentence in front of the microphone of the computer: "Happy New Year". See the Figure 2 above.

As shown on the Figure 2, the pronounced sentence was automatically recorded in the Search field of Google and the results of search appeared   in the main window of the program. The same principles apply to the members of the list, representing the search systems.

Consider the following possibility of the program. Let's mark the next category, "Google translate system" and pronounce desired sentence. See the Figure 3 below.

On the left side of the Figure 3, the shown sentence pronounced in front of the microphone: "The International scientific Conference" is recorded in the field of translation of the sentence, and on the right side of the same Figure, where is the field of translated sentence, the appropriate text pronounced in Georgian by a user, is shown in English language.
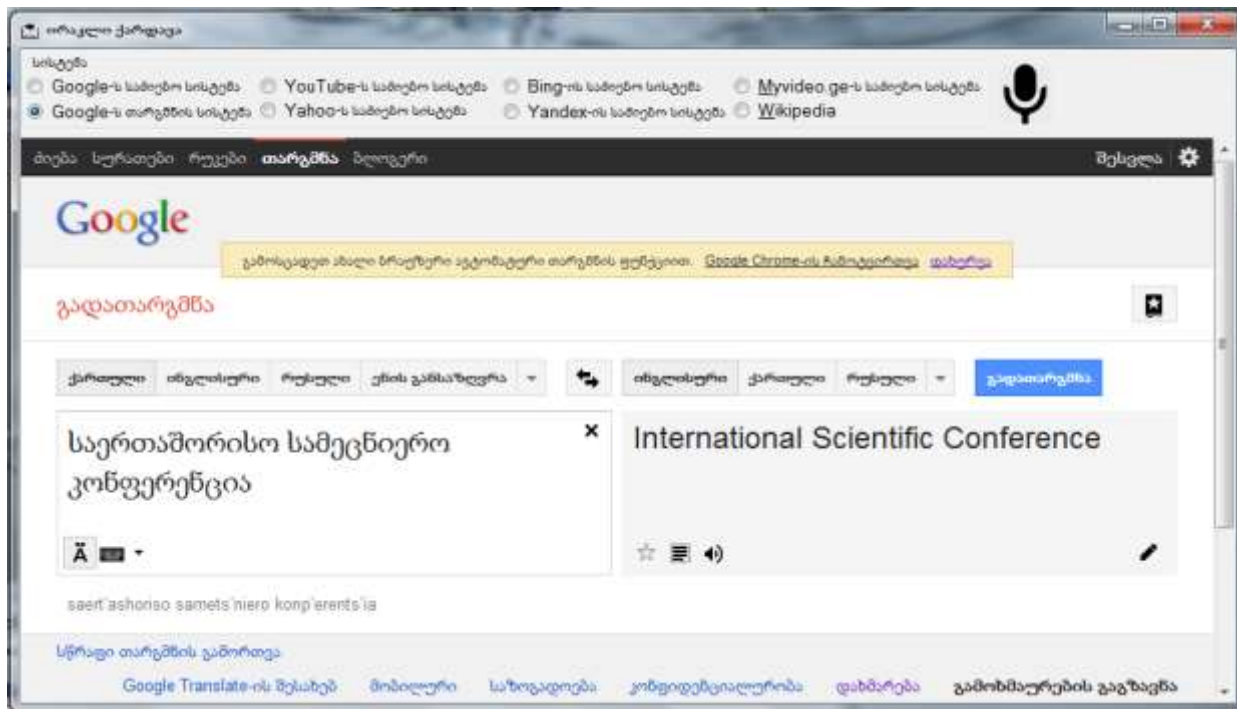
Figure 3. Google Translate system to translate Georgian pronounced sentences into English.

Besides, in concrete category of the program created by us, it is possible to choose the language, in which we want to translate the sentence pronounced in Georgian, without using the buttons, exactly as shown in the Figure 3 in the example of the English language.

Besides above mentioned issues, the program has the possibility to perform the characteristic functions of a web browser (Go back, Go forward, Minimize, Maximize, Refresh and so on.) with voice commands pronounced in Georgian language, without using any buttons. In case of existence of the word in the pronounced sentence composition matching the other voice command, the appropriate message window will appear, demanding to specify, in order, to avoid ambiguity, whether the pronounced sentence is command or part of the sentence. It should be noted, that this last function is performed in Georgian voice commands, see Figure 4.
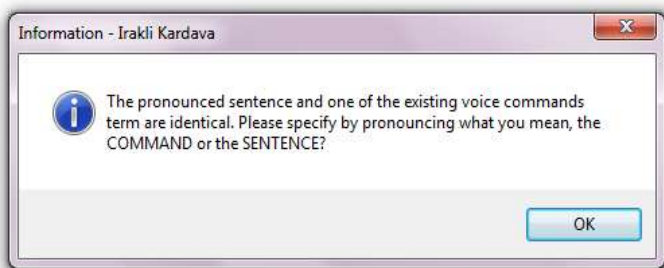

Figure 4. Information messagebox.

As it is shown in Figure 5, it is necessary and enough to pronounce "COMMAND" or "SENTENCE" for avoiding ambiguity and for further correctly work of the software. After choosing the desired variant, the program will perform following appropriate steps automatically.

The next possibility of the software is that the user can create his new voice0 commands or modify each of already existing voice commands, during working of the program.

In that case, when it is needed to create a new voice command, the user must pronounce pre-reserved special term "newcomm", after that the section of creating a new voice command will be activated and will be given a list of appropriate .NET Framework functions. After choosing a function, it is necessary to pronounce a word or a short sentence, which will be assigned to the desired function. If the process is finished successfully, the new voice command will be added to the old list of the commands. The same principle is used for modifying already existing voice commands. The difference is only in pre-reserved terms. In this case for activating section of modifying we have to pronounce "modcomm".

III.   The Scheme of Working of the Software

For illustrate of working of the software and its sub programs, below is given Figure 5, which shows detailed interdependence among each section of the whole process.
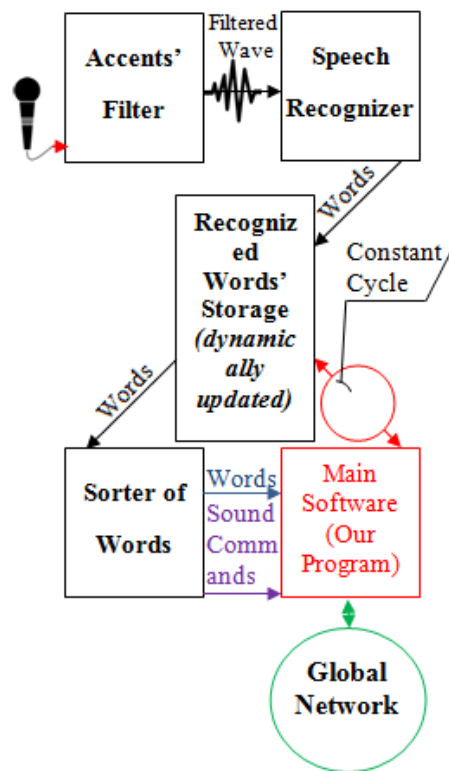
50

Figure 5. The Scheme of interdependence among each section of the whole program.

IV. **Conclusion**

By using the Software created by us, it is possible Global Network consumption and using the vast majority of its resources through the use of Georgian voice commands. Also, for automatic recording of Georgian pronounced words and sentences in modern and popular search or translate systems without using the keyboard. Development of this program will contribute the perfection of acoustic model of the Georgian language and integration in above mentioned systems, like other already realized languages (For example: English, Japanese, Italian, German, Spanish and etc.).

In general, by the realization of this approach, which is used for creation of our software, it is achievable to solve the artificial intelligence issues, such as arrange an improved sound dialogue between the computer (A machine) and human, independent from any accents of any languages. Not only for Georgian language, but for all languages, because the problem of the accents for speech recognition systems is a common problem.

The software is created in language C# by using of Microsoft Visual Studio 2008 Express Edition and Microsoft Visual Studio 2012. A Georgian speech recognition system is created by CMUSphinx. For sound surveillance, during solving the problem of the accents, we have used the system "Praat".

References

[1] N. Gamkrelidze, Sh. Kotetishvili, J. Lejava, L. Lortkipanidze, L. Javakhidze, *Phonetic Analysis of Georgian Normative and Dialectal Speech,* Tbilisi, Georgia: Nekeri, 2006.

[2] A.Shanidze. The Grammar of Georgian Language, Tbilisi, 1930 (in Georgian).

[3] D.MelikiShvili. The System of Georgian Verbs Conjugation, Logos Press, Tbilisi, 2001 (in Georgian).

[4] J.Antidze, N.Gulua, I.Kardava, The Software for Composition of Some Natural Languages' Words, Lecture Notes on Software Engineering, vol.1, #3, Singapure, pages 295-297, August 2013.

[5] Raj Reddy, *Speech Recognition by Machine: A Review*, Proceedings of the IEEE, 64, 501-531, April 1976.

[6] X. Huang, A. Acero, Hsiao-Wuen Hon, Spoken Language Processing: a Guide to Theory, Algorithm, and System Development, New Jersey: Prentice Hall PTR, 2001.

[7] D. Melikishvili, *The Georgian Verb: A Morphosyntactic Analysis*, New York, USA: Dunwoody Press, 2008.

[8] A. Black, K. Lenzo and V. Pagel, *Issues in Building General Letter to Sound Rules*. 3 rd ESCA Speech Synthesis Workshop, pp. 77-80, Jenolan Caves, Australia, 1998.

[9] T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi and T. Kitamura, *Simultaneous Modeling of Spectrum, Pitch and Duration in HMMBased Speech Synthesis,* Proc. of EUROSPEECH, vol.5, pp.2347–2350, 1999.

[10] K. Shinoda and T. Watanabe, *MDL-based context-dependent subword modeling for speech recognition,* J. Acoust. Soc. Jpn.(E), vol.21, no.2, pp.79–86, 2000.

[11] Fadi Biadsy, Automatic Dialect and Accent Recognition and its Application to Speech Recognition, Columbia University, 2011.

[12] Z. Wang, T. Schultz, and A. Waibel, "Comparison of acoustic model adaptation techniques on non-native speech", in ICASSP 2003. IEEE, 2003, pp. 540–543.

Irakli Kardava was born in Sokhumi, Georgia on Nov. 01, 1986. He graduated from Sokhumi State University with a master's degree in computer science, 2012.

He is a PhD student of computer science at The Sokhumi State University. He has 6 scientific publications. Irakli Kardava is a member of International Association of Computer Science and Information Technology (IACSIT, 9 Jurong Town Hall Road iHUB, Singapore 609431, www.iacsit.org).

From 2015 He is choosen for Erasmus Musnus – Eminence II, at Adam Mickiewicz University (Nanobiotechnology center) in poznan, poland.

For more information about him, please, visit the web site: http://www.irakli-geo.hol.es