Applications of Speech Recognition for Romanian Language

Cătălin CHIVU
"Transilvania" University of Braşov
str.Mihai Viteazu nr.5, RO-500174 Braşov
catalin.chivu@gmail.com

***Abstract*—In the world there were designed some speech recognition systems but none of these systems are dedicated to Romanian language. In the present paper are presented some specific applications for which is successfully applied a speech recognition system adapted by the author for Romanian language.**

***Index Terms* — computer programs, systems control, systems command, speech recognition.**

I. INTRODUCTION

In these days the *Speech Recognition* term is not a new one. Speech and voice recognition refers to the ability of machines to respond to spoken commands. Speech and voice recognition enables "hands-free" control of various electronic devices - a particular boon for many disabled persons - and the automatic creation of "print-ready" dictation. Among the earliest applications for speech & voice recognition was automated telephone systems and medical dictation software (for transcription).

The technology of Automatic Speech Recognition (ASR) and Transcription has progressed greatly over the past few years. Ever since research of this technology began in 1936, the largest barriers to the speed and accuracy of speech & voice recognition were computer speed and power (or the lack thereof). With the average CPU now above a Pentium IV and RAM levels at 1 GB and up, accuracy levels have reached 95% and better with transcription speeds at over 160 words per minute. There were developed, by various producers, many engines for speech recognition, with an increasing performance. These speech recognition systems may be divided in two categories:

1) speech recognition systems for *singular words or sentences* (implemented to low power computing devices), which support any language;

2) *continuous* speech recognition systems that are able to decompose long and very long phrases in phonemes and then recompose these phonemes in words and phrases according to a specific dictionary and a dedicated grammar. Such a speech recognition system is Dragon Naturally Speaking 9, developed by Nuance Company that, at this moment, according to the latest statistics, is the most used speech recognition system on the market. Another famous company that develops speech recognition systems is Microsoft. Microsoft developed the SAPI 5.1. pack destined to Windows users, which is a free pack of programs, well enough documented. The SAPI pack allows continuous speech recognition and automated reading (conversion of text to speech). Unfortunately, none of these speech recognition systems are dedicated for Romanian language.

However, the present paper presents a method used to adopt the most known speech recognition system, Microsoft SAPI 5.1 (designed for English language) to Romanian language.

Vocal recognition for Romanian language can be done by a vocal recognition system designed for this task. To be able to design such a system it should be defined a specific corpus that contains de training set of data, and the testing set of data [4]. Also, it should be create, starting from zero, the recognition system, than, this system should be trained and applied to particular tasks. The goal is to create a recognition system for continuous speech because, in this case, the system will be more applicable. The recognition engine is a program designed by the software companies, companies that invest fabulous sums of money in this field. The product of such a firm is a set of programs, executable ones. These are the reasons why the idea that one man can design such a recognition system is a utopia.

However, a specialist is able to design such a system by starting from an existing system, even if this one is already particularized for other language, and force it to function for Romanian language. This adaptation assumes to replace the grammar of recognition system (developed for English) with another grammar developed for the Romanian language and for a specific task. To rewrite the grammar specific to other language than the native one implies: redesign the desired expressions according to the system requirements; phonetic transcription of each Romanian word that appears in the current application, based on the native phonemic collection (in this case American English) [1].

II. VOCAL CONTROL APPLIED TO MANUALLY COMMANDED SYSTEMS

Manual command of the systems is useful for the human aided control and command of the industrial systems. For this type of command the system should be designed with a control and command switchboard where the human has access. The manual command is used on major scale even today, despite the high technology, because there are many situations that require a human interaction in decision process. An eloquent example is airplane driving. The small airplanes that require multiple and diverse actions to be driven have, especially manual control of flying parameters [5, 6].

In the following is presented an example of vocal control of a motor-valve system, manually commanded.

It was considered a very simple manually commanded system that has a dc motor that actuates a valve. The switchboard allows the human operator to see the pressure difference before and after the valve, the valve position or the flow rate of the fluid/air from the system.

Digital Object Identifier 10.4316/AECE.2007.01007

Electrical diagram of this system is presented in figure 1. This diagram is classical contact one and the purpose of this paper is to adapt the vocal control to this diagram.

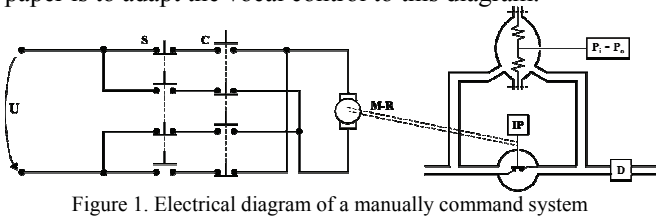


Figure 1. Electrical diagram of a manually command system

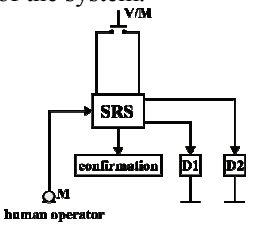
The significance of the notations from figure 1 is: *M-R* – actuator – reduction gear; *IP* – valve position indicator; P_i-P_o – indicator of the difference between the input and output pressure; *D* – flow rate indicator; *S* – decrease; *C* – increase.

The vocal command is applied to this system but it is maintained the manual command (the last one has priority when there is a contradictory vocal command). The vocal command implies the integration of a speech recognition system (SRS) in the command and control system of the device. This SRS is capable, depending on the vocal human commands, to rotate the motor in one way or the other with a specified step.

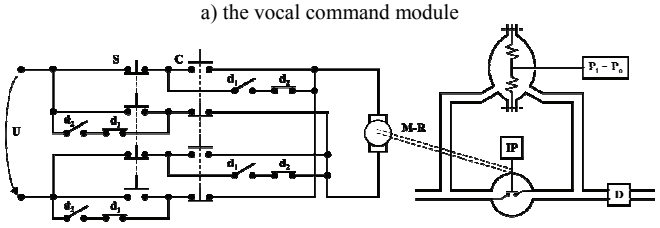
The diagram of this system is presented in figure 2. These diagrams can be implemented by using contacts or static commutation circuits.

In figure 2 the notations have the following significance: *Confirmare (Confirmation)* – module of visual or vocal confirmation of the received command; *M* – microphone; D_1, D_2 – command relays windings – increase and decrease; d_1, d_2 – contacts of D_1, D_2 relays; *SRS* – speech recognition system; *V/M* – control button for activation of vocal/manual command.

The system's ergonomics is significantly improved and depends on the quality of the speech recognition system. To create a high quality speech recognition system it should be created a grammar dedicated to this application. This grammar should define all the commands that could be given by the human operator. The vocal commands should substitute the manual ones and, eventually, to define a more accessible control of the system.



a) the vocal command module



b) the command electrical diagram

Figure 2. The manual-vocal command diagram of a manually commanded system

The human operator has to be able to do the following manually commands:

- to close the valve (the closing will be done proportionally to the time period during the human operator is pushing the close button);
- to open the valve (the opening will be done proportionally to the time period during the human operator is pushing the open button);
- to activate the vocal command (without inactivating the manual command);
- to partial inactivate the vocal command (without inactivating the manual one).

Using the vocal command the human operator can do the following tasks:

- close the valve (closing will be done with predefined steps or, as percentage of the valve stroke);
- open the valve (opening will be done with predefined steps or, as percentage of the valve stroke);
- partially inactivate the vocal command (without inactivating the manual command), the only active vocal command being the one of activating the vocal command;
- activate the vocal command (without inactivating the manual command), even if there is partially inactivated the vocal command;
- activate the vocal confirmation of a correct recognition of a word/expression by the speech recognition system;
- inactivate the vocal confirmation of a correct recognition of a word/expression by the speech recognition system.

Vocal commands have to implement the following valid operations:

- to close the valve, with a predefined step, the human operator has to say "*închide robinet*" ("*close valve*");
- to close the valve, with a percentage of the valve's full stroke, the human operator has to say "*închide robinetul cu 'x' procente*" ("*close valve with 'x' percentage*") or "*închide robinetul cu 'x' la sută*" ("*close valve with 'x' per cent*");
- to open the valve with a predefined step, the human operator has to say "*deschide robinet*" ("*open valve*");
- to open the valve, with a percentage of the valve's full stroke, the human operator has to say "*deschide robinetul cu 'x' procente*" ("*open valve with 'x' percentage*") or "*deschide robinetul cu 'x' la sută*" ("*open valve with 'x' per cent*");
- to fully activate the vocal command, the human operator has to say "*activează comanda vocală*" ("*activate vocal command*");
- to partially inactivate the vocal command, the human operator has to say "*dezactivează comanda vocală*" ("*inactivate vocal command*");
- to activate the confirmation, the human operator has to say "*activează confirmarea*" ("*activate confirmation*");
- to inactivate the confirmation, the human operator has to say "*dezactivează confirmarea*" ("*inactivate confirmation*");

In figure 3 is presented the interface of the application that implements the vocal command and control of the valve-electric motor system.

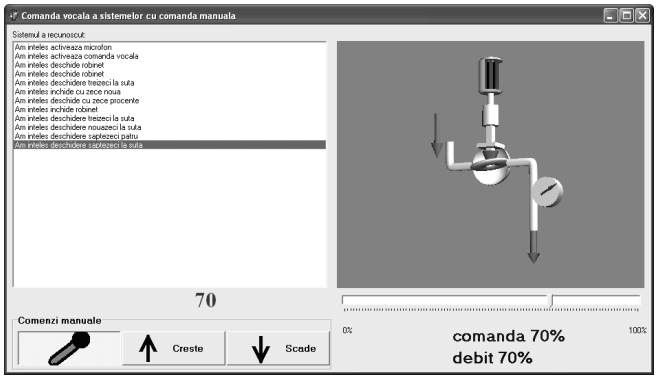


Figure 3. The interface of the program of vocal command and control

III. VOCAL COMMAND APPLIED TO AUTOMATIC CONTROLLED SLOWLY PROCESSES

The automatic controllers are based on a specific law and minimize the error of the system. Normally, an automatic controller has included the comparator, which compares the output with the input.

In the following it is considered a simple process controlled by an automatic controller that allows the operator to switch between the manual - automat operation modes. The process includes a D.C. motor that closes and opens a valve. The controlled parameter can be, depending on the situation, the pressure difference between input and output or the flow rate through the valve or the valve's position.

The vocal command module of such a system is shown in figure 4, and the command electric diagram is presented in figure 2b.

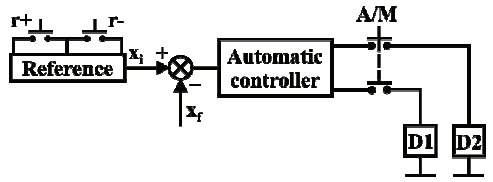


Figure 4. Diagram of vocal command module

The diagram from figure 4 commands the motor through the electric diagram presented in figure 2b.

The elements presented in the above figure represent:

Referință (Reference) – device that gives the reference value for the system;

r+ - button that increases the reference value;

r- - button that decreases the reference value;

x_i - reference value;

x_f - value of controlled parameter.

Regulator automat (automatic controller)

System presented in figure 4 allows both the manual and automat control of the process. The human operator can use manual command to operate the system.

The scheme from figure 4 allows to control the system both manually and automatic. The human operator is able to use, at any moment, the manually control to command the process. If the system is working in automatic operation mode and the human operator use manual commands then, temporarily, the process switch on the manual command. In this situation, when the manual command stops, the system switches back on automatic operation mode. The automatic controller may be manually activated or inactivated by activating or inactivating the command on the D_1 and D_2 relays. If the automatic controller was inactivated the process control is a manual one obtained by using the *Scade (Decrease)* and *Creşte (Increase)* buttons (figure 3).

Switching on between manually and automatic command is done without any problem because the motor position (and the valve position), which is steady during no supply time range, memorizes the output variable until is switched off the current command.

The reference can be a constant or manually modified by the human operator.

The purpose is to apply vocal command to such a system by maintaining active the switching possibility between automatic and manually control. This condition implies to include the Speech Recognition System (SRS) in the command and control system. Depending on the type of command that it receives (manual or vocal), SRS is able to command the switch on/off between manually and automatic control, to rotate the motor in one direction or the other with a specific angle, if the command is manually, or, if the command is an automatic one, to change the reference of the controller. The reference can be changed both manually and vocal.

The vocal command module of such a system is shown in figure 5 and the electric diagram was presented in figure 2b.

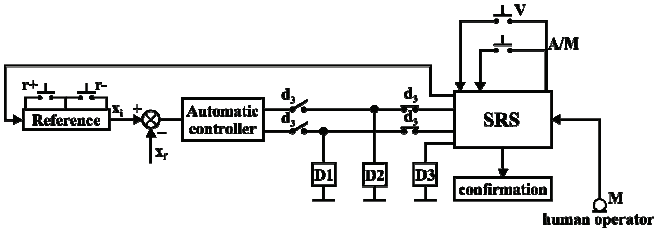


Figure 5. Vocal command module of a discrete automatic control system

Using scheme presented in figure 5 there can be transmitted commands to the motor through an electric circuit that has the diagram presented in figure 2b.

The notations from the above figure represent:

D_3 - relay winding (the relay command the switching on between Automatic and Manually operation mode),

d_3 - contacts of D_3 relay,

V - button that allows to activate or inactivate the vocal command,

A/M - button that switches on/off the Automatic – Manually operation mode.

Speech recognition system, used as part of the automatic control system, improves the ergonomics of the system. Also, it is necessary to create a grammar dedicated to this application. This grammar has to include all the possible forms for the commands transmitted by the human operator to the system. The vocal commands have to completely substitute the manual commands and, eventually to facilitate to system control.

In the case of manually command, the human operator has to be able to do the following manually operations:

- close the valve (proportionally to the time interval during which the close button is hold on);
- open the valve (proportionally to the time interval during which the open button is hold on);
- activate the vocal command (without inactivating the manual one);
- partially inactivate the vocal command (without inactivating the manual one);

- modify the reference;
- switch on the automatic command.

In the case of automatic command the human operator has to be able to do the following manually operations:

- activate the vocal command;
- partially inactivate the vocal command;
- choose the reference;
- switch on the manually command.

For the manually command, the human operator has to be able to do the following vocal commands:

- close the valve (with predefined steps or percentage of the valve's full stroke);
- open the valve (with predefined steps or percentage of the valve's full stroke);
- switch on the vocal command (without inactivating the manual one), even if the vocal command is partially inactivated;
- partially inactivate the vocal command (without inactivating the manual one);
- change the reference (by increasing or decreasing with a predefined step or percentage);
- switch on the automatic command;
- activate the vocal confirmation of a correct recognition done by the Speech Recognition System;
- inactivate the vocal confirmation of a correct recognition done by the Speech Recognition System.

For the automatic command, the human operator has to be able to do the following vocal commands:

- activate the vocal command (without inactivating the automatic command), even if the vocal command is partially inactivated;
- partially inactivate the vocal command;
- choose the reference (by increasing or decreasing with a predefined step or percentage of the definition interval);
- switch on the manually command;
- activate the vocal confirmation of a correct recognition done by the Speech Recognition System;
- inactivate the vocal confirmation of a correct recognition done by the Speech Recognition System.

The vocal commands have to be part of the following possible operations:

- to close the valve with a predefined step the human operator has to say "*închide robinet*" ("*close valve*");
- to close the valve with percentage of the valve's full stroke, the human operator has to say "*închide robinetul cu 'x' procente*" ("*close valve with 'x' percentage*");
- to open the valve with a predefined step the human operator has to say "*deschide robinet*" ("*open valve*");
- to open the valve with percentage of the valve's full stroke, the human operator has to say "*deschide robinetul cu 'x' procente*" ("*open valve with 'x' percentage*");
- to activate the vocal command, the human operator has to say "*activează comanda vocală*" ("*activate vocal command*");
- to inactivate the vocal command, the human operator has to say "*dezactivează comanda vocală*" ("*inactivate vocal command*");
- to activate the confirmation, the human operator has to say "*activează confirmarea*" ("*activate confirmation*");
- to inactivate the confirmation, the human operator has to say "*dezactivează confirmarea*" ("*inactivate confirmation*");
- to change the reference, by increasing the value with a predefined step, the human operator has to say "*creşte referinţă*" ("*increase reference*");
- to change the reference, by decreasing the value with a predefined step, the human operator has to say "*scade referinţă*" ("*decrease reference*") or "*decrementează referinţă*" ("*decrement reference*");
- to change the reference to a specific value, the human operator has to say "*mută referinţa la 'x' la sută*" ("*change reference to 'x' per cent*") or "*mută referinţa la 'x' procente*" ("*change reference to 'x' percentage*");
- to switch on to automatic command, the human operator has to say "*treci pe comandă automată*" ("*switch to automatic command*");
- to switch on to manually command, the human operator has to say "*treci pe comandă manuală*" ("*switch to manual command*").

In figure 6 is presented the interface of the program that implements the speech recognition to the command and control of a system with automatic control and that allows to switch on between the Automatic and Manual operation mode.

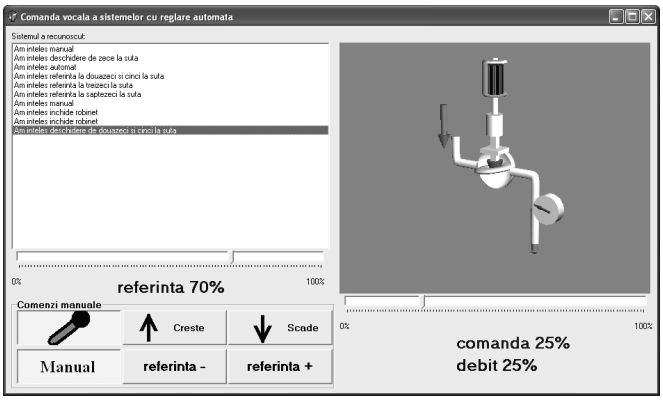


Figure 6. Interface of the program that applies vocal command to an automatic control system

IV. CONCLUSION

Speech recognition may by successfully apply to manually command system, when the human operator works in a noiseless environment.

The speech recognition system increases the ergonomics of a work space because there can be defined commands, with variable complexity, and assigned these commands to many operations (that are also available manually).

The advantages of this speech recognition are useful to control complex systems, when, even if the human operator has to command manually the system, he can also use this "assistant" to do his task. There are many applications when the commands are multiple and complex and should be done synchronized, in this case one operator could not command the system just manually. Using a speech recognition system could be defined commands that actuate sets of switchers, at arbitrary moments, and so, the system could be controlled only by one human operator.

The disadvantages of the speech recognition system appear when the system is very complex, which implies a

high degree of difficulty for the vocal commands and human operator may confound the vocal expressions. Another disadvantageous situation is the case when the vocal commands, assigned to particularly activities, are almost the same or are very similarly, case when the speech recognition system constantly confounds them. To solve these problems the grammar of the system should be very well structured, thus the commands will be as simple as possible and the used words combinations should eliminate the ambiguities.

In last application the automatic control commands and control, most of the time, the system. The manual command is useful, especially, when appears an abnormal function (when the system starts and stops, when, different type of loads are connected or disconnected). When this malfunctioning appears the human operator can change, manually, the reference of the automatic controlled system or can switch on the manual control, thus the whole control of the system is done manually.

The vocal command is very useful when human operator has to accomplish many tasks in the same time, or sequential, but in short periods of time. Using this vocal command the human operator is able to modify working parameters of a process during he is changing, manually, other parameters of the same process or of other process.

REFERENCES

[1] Chivu, C.I, Contribuții privind procesarea semnalelor vocale (Contribution regarding signal processing). PhD Thesis. Politehnica University of Bucharest.

[2] Dumitrache, I., Constanton, N., Drăgoicea, M, - Rețele Neurale. Identificarea şi Conducerea Proceselor. Editura MATRIX ROM, Bucureşti, 1999.

[3] Ganciu, T., Horga V. - Identificarea sistemelor, Ediția a II-a revizuită, Editura "Gh. Asachi" Iaşi 2002.

[4] Giurgiu, M. - Results on Automatic Speech Recognition in Romanian. Recent Advances in Romanian Language Technology, Ed. Academiei Române, 1997.

[5] Nitu, C., Matlac, I., Feştilă, C., - Echipamente electrice şi electronice de automatizare. Editura didactică şi pedagogică, Bucureşti, 1980.

[6] Popescu, D., Lupu, C., Petrescu, C., Alexandru, M., Mateescu, M. – Sisteme de Conducere a Proceselor Industriale. Editura Printech, Bucureşti, ISBN 973-718-016, 2004.