ANNALS of the ORADEA UNIVERSITY.
Fascicle of Management and Technological Engineering, Volume VI (XVI), 2007

ROMANIAN CONTINUOUS SPEECH RECOGNITION APPLIED TO AUTOMATIC CONTROLLED SYSTEMS

Cătălin Chivu
"Transilvania" University of Braşov, catalin.chivu@gmail.com

Keywords: speech recognition, vocal command, grammar, controller

Abstract: The continuous speech recognition applied to Romanian language is field not very known for the specialists. There are many recognition systems (engines) developed by the large companies but these systems are designed for other languages than Romanian. This paper presents a grammar created for a specific application (vocal command of an automatically controlled system), grammar that is based on an adaptation of a recognition engine, designed for English language, to Romanian language.

1. INTRODUCTION

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. Also, it should be created, 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.

Specialists are 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. To adapt the system it should be replaced the grammar of the recognition system (the grammar is designed for applications in English) with other grammar, special designed for the desired application in Romanian language. To rewrite the grammar of the native language of the system for other language, different of the native one, implies, among redesigning the desired expression, according to the rules assess by the system, the phonetic transcription of each Romanian word, which is part of the new grammar, based on the phonemic collection of the native language of the recognition system (in this situation the American English language).

First of all, a recognition engine, to be able to be personalized (to recognize words from specific languages), must have defined the following elements:

- numbers of phoneme and the code associated with these phonemes; these correspond to the specific language;
- the links between phonemes;
- the set of words that will be recognized and that belongs to the specific language;
- at least one elementary grammar, which defines a relation between the words that are part of the set that will be recognized.

The words included in the set that will be recognized should be defined based on the phonemes specific to that language.

In this paper it is presented the grammar specific to the vocal command of a system that is automatically controlled.

2. PHONEMIC REPRESENTATION OF ROMANIAN LANGUAGE

Although English language is not a phonetic one, most of the words are pronounced totally different that are written, the phonetic written rules of English language directly connect the pronunciation of the words with the written transcription of them [4].

If there is analyzed the phonemic collection of the SAPI5 toolbox, it can be observed a certain similarity with the phonemic collection of Romanian language.

The *aa* phoneme from the English word *father* is the same with that from the Romanian word *ladă*. This conclusion can be demonstrated by comparing the two spectrograms: first of the English word *father* (figure 1), second of the Romanian word *ladă* (figure 2) [1].

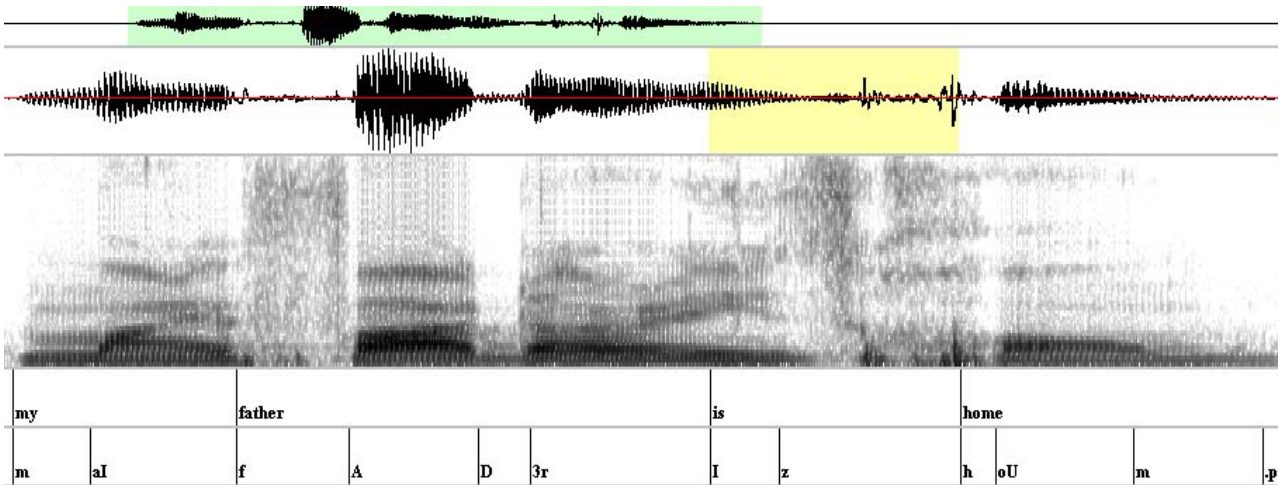


Figure 1 Phonemic label for expression "my father is at home"

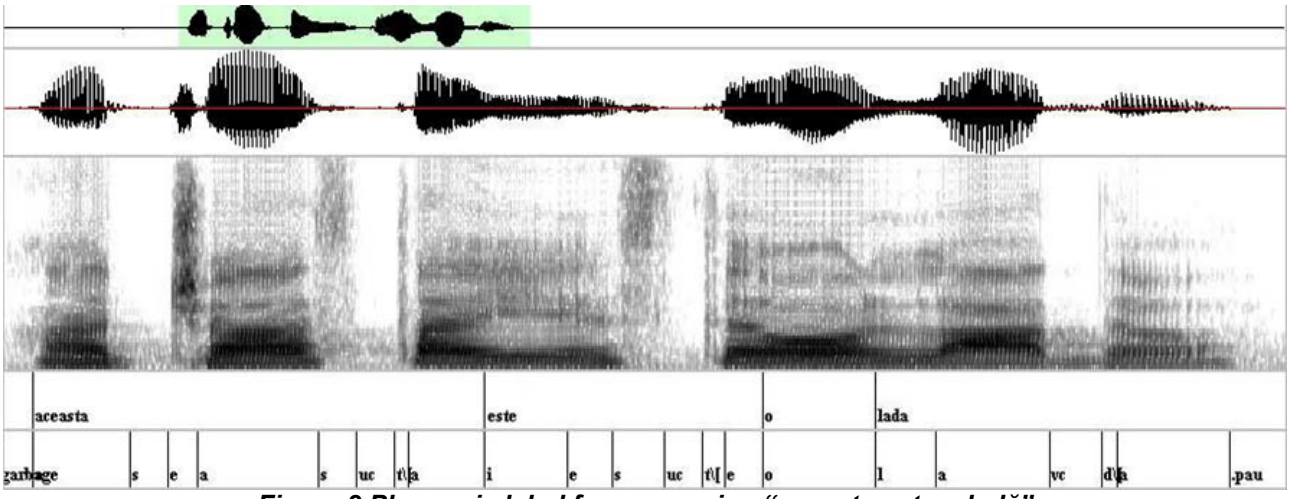


Figure 2 Phonemic label for expression "aceasta este o ladă"

In the following there is presented the correspondence between English and Romanian phonemes [1].

English phoneme	English word	Romanian word
ax	*ago*	*hău*
b	*big*	*stabil*
ch	*chin*	*cină*
d	*dig*	*dig*
eh	*pet*	*pet*
er	*fur*	*măr*
f	*fork*	*foc*
g	*gut*	*gaz*
h	*help*	*han*

English phoneme	English word	Romanian word
ih	*fill*	*stil*
jh	*joy*	*imagine*
k	*cut*	*cal*
l	*lid*	*lin*
m	*mat*	*mal*
n	*no*	*năr**av*
ow	*go*	*nod*
p	*put*	*puț*
r	*red*	*rest*
s	*sit*	*dosit*
sh	*she*	*şiret*
t	*talk*	*talc*
uh	*book*	*suc*
v	*vat*	*vară*
z	*zap*	*zar*
zh	*pleasure*	*jar*

However, it can be observed that the base phoneme *î* (i.e. from Romanian word *cât*) is missing.

That is way this phoneme will be substitute with *ax* phoneme, presented above, because this one has the closest pronunciation. The other Romanian phoneme that has no correspondent can be described using combination of two or more English phonemes.

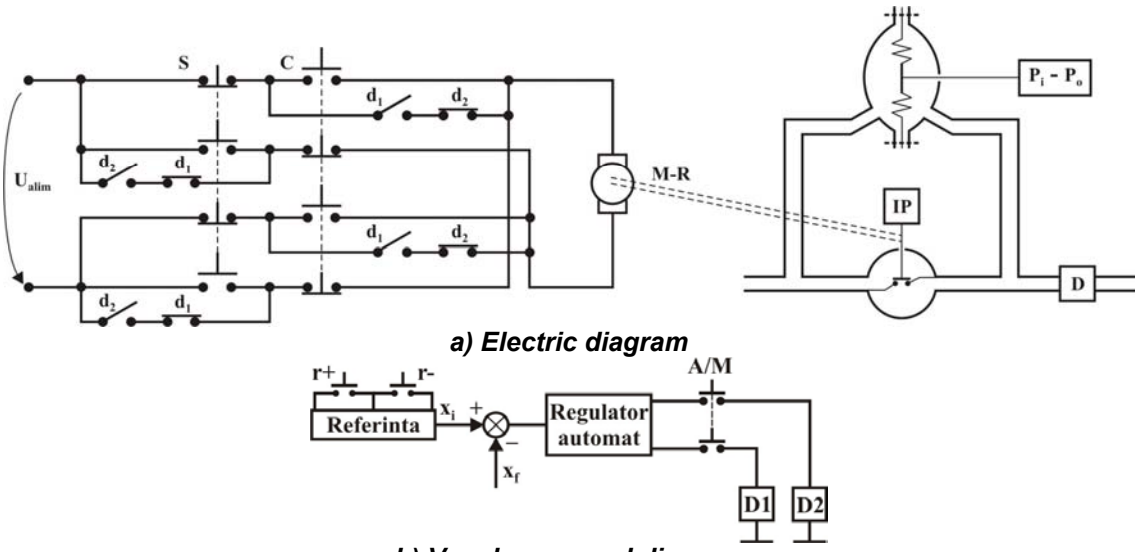
The *ch eh* phonemes correspond with the phoneme for the group *ce,* from word *cenuşã*. The *k eh* phonemes correspond with the letters group *che* from *cheie* word. The *k ih* phonemes correspond with the letters group *chi* from *schit* word. The *d jh ae* phonemes correspond with the letters group *ge* from *gel* word. The *g y eh* phonemes correspond with the letters group *ghe* from *ghem* word. The *d jh ih* phonemes correspond with the letters group *gi* from *gin* word. The *g ih* phonemes correspond with the letters group *ghi* from *ghid* word. The *t z* phonemes correspond with the letters group *ț* from *țap* word [1].

3. 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 [2, 5, 6].

In the following it is considered a simple process controlled by an automatic controller that allows the operator to switch between the manual and 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 command electric diagram is presented in figure 3a and the vocal command module of such a system is shown in figure 3b.



a) Electric diagram

b) Vocal command diagram

Figure 3. Diagram of vocal command module [1]

The diagram from figure 3b commands the motor through the electric diagram presented in figure 3a.

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)

The scheme from figure 3 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 uses manual commands then, temporarily, the process switches 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 5).

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) [1, 4] 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 4 and the electric diagram was presented in figure 3a.

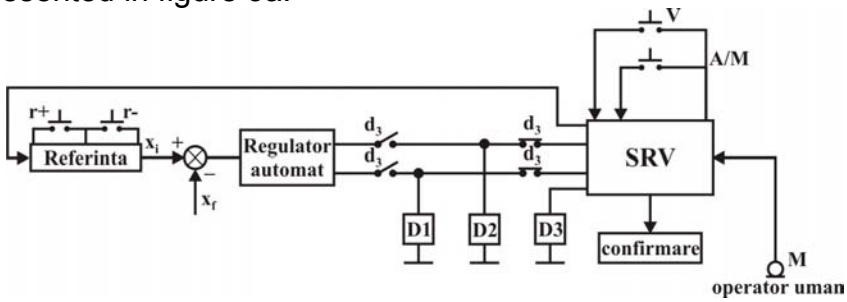


Figure 4. Vocal command module of a discrete automatic control system

Using scheme presented in figure 4 there can be transmitted commands to the motor through an electric circuit that has the diagram presented in figure 3a.

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 the 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*") 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 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 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 referință*" ("*increase reference*") or "*incrementează referință*" ("*increment reference*");
- to change the reference, by decreasing the value with a predefined step, the human operator has to say "*scade referință*" ("*decrease reference*") or "*decrementează referință*" ("*decrement reference*");
- to change the reference to a specific value, the human operator has to say "*mută referința la 'x' la sută*" ("*change reference to 'x' per cent*") or "*mută referința 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 5 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.

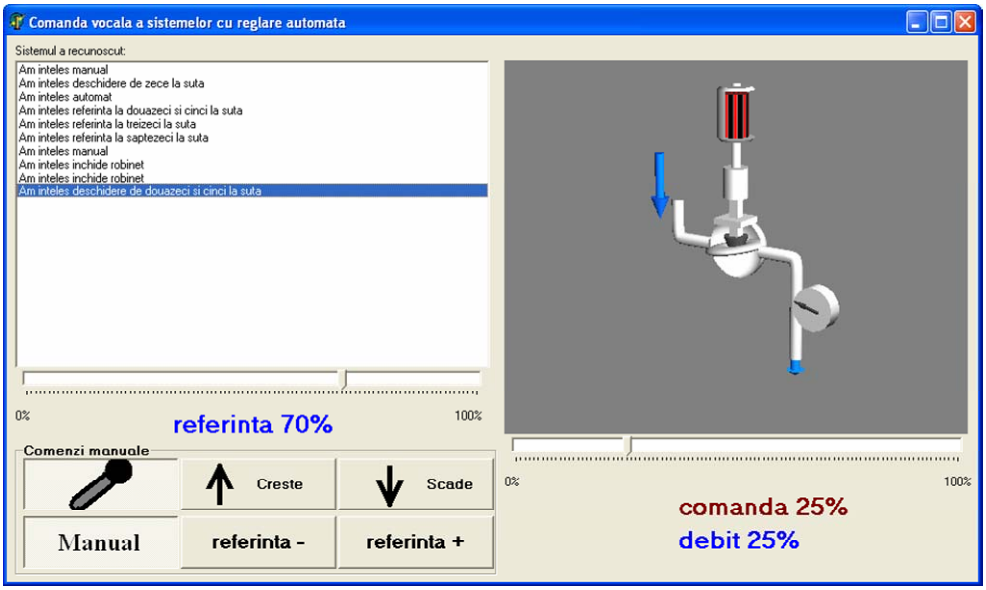


Figure 5. Interface of the program that applies vocal command to an automatic control system [1]

4. GRAMMAR DEDICATED TO VOCAL COMMAND OF THE SYSTEM

As it was specified above, to be able to implement a vocal command, it has to be created a grammar dedicated to the specific application.

In the following is presented a short sequence of the grammar dedicated to the application presented above in this paper [1].

- <GRAMMAR LANGID="**409**">
- <DEFINE>
<ID NAME="**PID_comenzi_rapide**" VAL="**31**" />
…
</DEFINE>
- <RULE NAME="**rscurte**">
- <L PROPNAME="**scurte**" PROPID="**PID_scurte**">
<P VAL="**1**">**/inchide aplicatia/inchide aplicatia/ax n k ih d ae aa 1 p l iy k aa 1 t z iy aa 1 ;**</P>
<P VAL="**2**">**/inchide robinet/inchide robinet/ax n k ih d eh r ow b iy n eh t;**</P>
<P VAL="**3**">**/deschide robinet/deschide robinet/d eh s k ih d eh r ow b iy n eh t;**</P>
<P VAL="**4**">**/activeaza comanda vocala/activeaza comanda vocala/aa 1 k t iy v eh aa 1 z aa 1 k ow m aa 1 n d aa 1 v ow k aa 1 l aa 1;**</P>
<P VAL="**4**">**/activeaza microfon/activeaza microfon/aa 1 k t iy v eh aa 1 z aa 1 m iy k r ow f ow n;**</P>
<P VAL="**5**">**/dezactiveaza comanda vocala/dezactiveaza comanda vocala/d eh z aa 1 k t iy v eh aa 1 z aa 1 k ow m aa 1 n d aa 1 v ow k aa 1 l aa 1;**</P>
<P VAL="**5**">**/dezactiveaza microfon/dezactiveaza microfon/d eh z aa 1 k t iy v eh aa 1 z aa 1 m iy k r ow f ow n;**</P>
<P VAL="**5**">**/inactiveaza comanda vocala/inactiveaza comanda vocala/iy n aa 1 k t iy v eh aa 1 z aa 1 k ow m aa 1 n d aa 1 v ow k aa 1 l aa 1;**</P>
<P VAL="**5**">**/inactiveaza microfon/inactiveaza microfon/iy n aa 1 k t iy v eh aa 1 z aa 1 m iy k r ow f ow n;**</P>
<P VAL="**6**">**/activeaza comanda manuala/activeaza comanda manuala/aa 1 k t iy v eh aa 1 z aa 1 k ow m aa 1 n d aa 1 m aa 1 n uw 1 aa 1 l ax 1;**</P>

```
  <P VAL="6">/treci pe comanda manuala/treci pe comanda manuala/t r eh ch p eh k ow m aa 1 n d ax
1 m aa 1 n uw 1 aa 1 l ax 1;</P>
  <P VAL="6">/comanda manuala/comanda manuala/k ow m aa 1 n d ax 1 m aa 1 n uw 1 aa 1 l ax
1;</P>
  <P VAL="6">/manual/manual/m aa 1 n uw 1 aa 1 l ;</P>
  <P VAL="7">/activeaza comanda automata/activeaza comanda automata/aa 1 k t iy v eh aa 1 z ax 1 k
ow m aa 1 n d aa 1 uw 1 t ow m aa 1 t ax 1;</P>
  <P VAL="7">/treci pe comanda automata/treci pe comanda automata/t r eh ch p eh k ow m aa 1 n d
aa 1 uw 1 t ow m aa 1 t ax 1;</P>
  <P VAL="7">/comanda automata/comanda automata/k ow m aa 1 n d aa 1 uw 1 t ow m aa 1 t ax
1;</P>
  <P VAL="7">/automat/automat/aa 1 uw 1 t ow m aa 1 t ;</P>
  <P VAL="8">/creste referinta/creste referinta/k r eh sh t eh r eh f eh r iy n t z ax 1;</P>
  <P VAL="9">/scade referinta/scade referinta/s k aa 1 d eh r eh f eh r iy n t z ax 1;</P>
  </L>
  </RULE>
- <RULE NAME="rcompuse" TOPLEVEL="INACTIVE">
- <L PROPNAME="compuse" PROPID="PID_compuse">
  <P VAL="1">/deschidere/deschidere/d eh s k ih d eh r eh;</P>
  <P VAL="2">/deschide cu/deschide cu/d eh s k ih d eh k uw 1 ;</P>
  <P VAL="3">/inchide cu/inchide cu/ax n k ih d eh k uw 1;</P>
  <P VAL="4">/referinta la/referinta la/r eh f eh r iy n t z ax 1 l aa 1;</P>
  <P VAL="5">/creste referinta cu/creste referinta cu/k r eh sh t eh r eh f eh r iy n t z aa 1 k uw 1 ;</P>
  <P VAL="6">/scade referinta cu/scade referinta cu/s k aa 1 d eh r eh f eh r iy n t z aa 1 k uw 1;</P>
  </L>
  </RULE>
```

5. CONCLUSION

The phonemic correspondence is successfully working because the speech recognition system decomposes the vocal signal in frames, gives probabilities to each phoneme for each frame and then, the Viterbi search selects the words that have the highest total probability. A hyphened phoneme of Romanian language will be automatic decomposed in phonemes specific to speech recognition system, on the same time when the signal is decomposed in frames. If the Romanian phoneme is described as precise as possible, using the speech recognition system, the phoneme may obtain a high probabilistic score and so, the word will be recognized.

Speech recognition may by successfully applied 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).

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.