

An Electronic Game with Speech Recognition for Children with Problems of Speech

Carlos Miranda-Palma, Reyna Camal-Uc, José Cen-Magaña, Cinhtia Gonzalez-Segura, Sergio Gonzalez-Segura, Manuel Escalante and Jorge Ríos

Universidad Autónoma de Yucatán, Facultad de Matemáticas, Unidad Tizimin, Calle 48 A s/n x 31, 97700, Tizimin, Yucatán, México

cmiranda@uady.mx, {camaluc, cmagana}@tizimin.uady.mx, {gsegura, sgsegura, etorres, rmatine}@uady.mx.

Abstract. In this work, we describe an electronic game (Memory Game) with speech recognition for stimulation on children at scholar age with problems of speech. We present the characteristics and elements of electronic game, as well as its integration with speech recognition system. For developed of the electronic game we used the Language C and the SDL graphic libraries and for the acoustic models we used the HTK software of speech recognition.

Keywords: Electronic game, speech recognition, children, problems of speech.

1 Introduction

Speaking is a tool that human being uses to exchange information. However, there are people with speaking disorders, and it is difficult or impossible to them to communicate in this way.

A speech and language disorder refers to the problems of communication or other areas related to it, for instance the motor oral functions. These disorders vary from simple substitutions of sound to the disability to understand or use language to speak or feed [1].

The communication disorders have the potential of isolate to the individuals in their social and educational environment, and because of that it is necessary to find out mechanisms to increase life-quality of people with special needs, so they can contribute to society.

Technology can help to those children whose physical conditions make the communication difficult. The use of electronic systems of communication let people who don't talk or have severe disabilities in speech increase their participation [1].

Because of the importance of speaking in people's life, the development of systems able to accept a voice signal as input to do a determinated task has become a big challenge. Actually, great quantities of systems which use human voice are being developed, the ITESM Campus Cuernavaca that by means of speech commands a mobile robot is manipulated [2][3]. At IIMAS, from UNAM, had been developed several projects that included speech communication as a interaction form [4][5].

In the same way, in the latest years had been incremented the development of applications for people with some kind of disability, For example the Projecte Fressa [6] that had development successfully several applications for different kinds of disabilities. One of this programs is the "Mouse Control through voice commands" that allows to control the mouse by means of the voice, it was designed for people that can not control the mouse in a regular form and they could have the control of their voice [6].

In this work we presented an electronic game (memory game) that used the most natural way of human communication: the Speech. This game is designed to be used with children in scholar age that have speech problems. Its characteristics, its environment and its integration with the recognition of voice system are described.

2 Memory Game

In this section we presented the conceptual design of Fonetix system and the memory game, as well as the elements and general characteristics of the game, its history, the interfaces and its characters.

2.1 Conceptual Design

When this project began, we proposed develop applications that were interesting, motivating and that support to children with language problems, for this reason and working co-ordinately with teachers specialized in language area from the Multiple Attention Center (MAC), located in Tizimín, Yucatan, we decided to implement this game, mainly because children enjoy playing it. Once the game had been defined, we didn't want this application was just the game itself, but we wanted to develop a story too, with an environment known by children from this region of the country. Furthermore, it was decided (with the teachers from MAC) to use a collection of direct and indirect words to speech recognition. A direct word is formed by syllables which begin with a consonant, followed by a vowel. An indirect word is formed by at least one syllable that begins with a vowel, followed by a consonant.

The next step we decided that all game story is developed in a house and that the environment from the memory game was a room of the same house. In the same way 4 characters were defined among them, the child could choose and this character the child helps to the character to defeat the challenges that an Alux (mythological character of Mayan culture) presents him during the interaction of the game.

In the Mayan mythology, an Alux is a little spirit that travels through the earth. An Alux is a little taller than an adult person's knee. It wears traditional Mayan clothes; generally is invisible but become into physical form to communicate with human beings.

From the previous, Fonetix system was born. In its first stage the system contains the memorama game. This system recognizes direct and indirect words from child's voice. Actually we are developing two more games, with other collections of words that are used in children's rehabilitation.

2.2 History of Game

At the beginning of the game, an introduction is presented the character that you could chose (Rafa, Tony, Luci or Migue) they have their toys and another articles disordered for all the room, in this moment their mother enters and warns them that if he/she does not order his/her room the Alux will come and take out all his/her things. The character does not obey his/her mom, maybe because he/she does not believe in the Aluxes or maybe because is very lazy, but at the moment the character decides to go out of his/her room but the Alux appears. Seeing the situation, the Alux decides to have fun with the child for a while, and turns every thing in the room into stone. These "stones" fly in the air, disappear and appear like stones in a huacal that is in the room.

Huacal is a term in Mayan language that means a type of cage or box, formed by separated loads of wood, and it is used to transport fruits, vegetables or small animals. After that the Alux informs to the character that if he/she touch 2 stones and them results being objects from the huacal he/she has to pronounce the name of the object, if he/she pronounce correctly the name of the object, this are going to be ordered automatically, but he/she only has 3 chances to pronounce the name of the object. The Alux disappear and leaves the character sad and with a lot of stones inside the huacal. Now is the turn of the child to find the objects

2.3 Description of the Game

Memorama is a game in which the child has to find the pair of stones that contains the same object, touching only two of them in each opportunity. This will let the child to put every object in the correct place in the room. The game has basically two views. The first view is the memorama game itself, represented by the set of stones in the huacal. The another view is the character's room, which contains the untidy objects first, and the animation of the found objects at last.

The game begins with the selection of one of the four characters which the child will use. After that, the Alux's inscription is presented, which turns the object into stones. Once this part has finished, it will be presented to the child the view of the room. This view will be different for each character (Fig. 1). In the upper corner of the view is presented the image of the selected character, as well as the huacal that contains the objects become into stones. The huacal is the link to go to the view of the memorama game.

When you select the image from the huacal, is presented the screen that represent the bottom of the huacal with the stones inside, in the superior part is presented the score of the game (Fig. 2).

To select a stone, the child locates and selects it with a click of the mouse. In this way when the player select a stone this turns in its original form with the name of the object in the lower part, and stayed in this ways until the player select a second stone (Fig. 2). If at any moment, the player wants to see the view of the room, he only selects the small house located in the upper part of the current view. The player can exchange between the two views the times and the moment he wants.

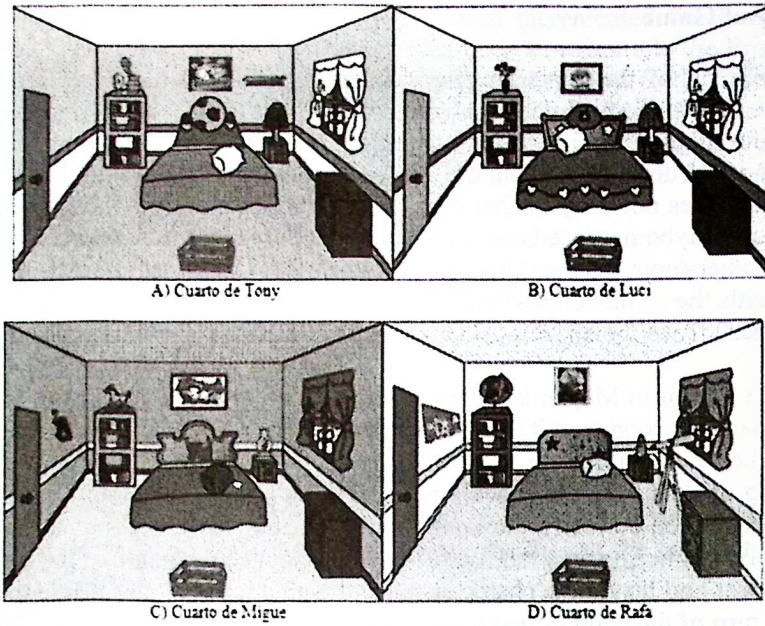


Fig. 1. Images Tony, Luci, Migue and Rafa's rooms.

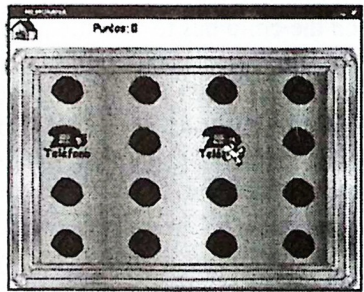


Fig. 2. View that represents the bottom of the huacal, with the stones inside.

When the player selects a pair of stones, two situations are presented: the first is when the stones selected contain different objects. Then, the objects turn into stones again, hiding their form and the game continues. The second case is when the objects in the stones are the same. When this happens, the system asks the user to pronounce the name of the object correctly, giving him three opportunities to find it. If the option of sound is abled, the system pronounces the word before the user can do it. It is important to say that this option is very useful for a child who doesn't read or those who need to exercise their audition.

If the player can pronounce correctly the name of the object, then the system gives him points: If the word was pronounced correctly in the first opportunity, the system gives 400 points. If it was obtained in the second chance, they are 300 points; in the

last try, only 200 points are regarded. Besides the points obtained, the player can see in the view of the room an animation of the pronounced object (Fig. 3).

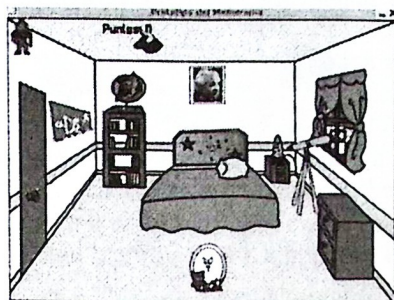


Fig 3. Animation of the correctly pronounced object.

If in the three opportunities the player can't pronounce correctly the name of the object, the system just gives him 100 points and the animation of the object is not presented.

When the objects are the same, the pair of stones selected disappears from the screen. The game ends when all the objects have been found; that means there are no more stones in the huacal and all the objects are in the correct place in the room. The system shows the room tidied up.

Once the game has finished, a chart is presented to the player, which contains the punctuations, and shows the ten best positions and their times. So, the goal of the game is to tidy the room up, turning the stones into objects, obtaining the better punctuation in the less possible time.

Finally, the player sees a screen where the system gives the opportunity to play the next level or exit the game. It is important to mention that the game has three levels of complex (number of direct and indirect words to be pronounced): in the first level there are 4 words, in the second level there are 8 words, and in the third level there are 12 words.

2.5 Characters

In this section we will present the characters of the game. Their characteristics were designed according to the children's environment, so they are identified with each character. It is important to mention that the names of the characters was put to the vote and were chosen by children that assist to Multiple Attention Center, from Tizimin.

The first of them is Tony (The sportsman): He is 9 years old and he likes a lot sports, he likes to works in teams and he prefer the sports in group, because he has the opportunity to spend time with his friends. His favourite food is stew (Fig. 4).

The second of them is Rafa (The studios): He is 10 years old, he likes reading and he help his partners with their homework. His favorite subjects are math and astronomy, because when he grows up he wants to be an astronaut. His hobbies are: seeing the stars with his telescope and solving riddles (Fig. 5a).



Fig. 4. Tony



Fig. 5. a) Rafa, b) Luci, c) Migue's images

The third of them is Luci (The romantic): she is 8 years old, she likes listening to music and reading poems. lessons and she loves helping her mother baking cookies (Fig. 5b). The last of them is Migue (The naughty): he is Luci's younger brother, he is 6 years old, he likes cartoons and collecting insects. He wants to be a cowboy, like his father. On Sundays, he likes go to the market to eat 2 cochinita tacos and drink horchata (Fig. 5c).

3 Speech Recognition System

The speech recognizing system used 3 main models:

- A *language model* which is based on a finite-states grammar. In this grammar there are defined all the words, direct and indirect, used in the memorama game. With this grammar, 60 words are recognized. These words were defined co-ordinately with the teachers of CAM. Some examples are: *ardilla, aspiradora, espada, escoba, vaso, zapato, dado, corona, cepillo, basura, espejo, hamaca, lupa, goma, galleta, payaso, bate, jirafa, camisa, antifaz, ovni, etc* (These Spanish words are means in English is: *squirrel, vacuum cleaner, spade, broom, glass, shoe, dice, crown, brush, trash, mirror, hammock, gum, cookie, clown, bat, giraffe, shirt, mask, ufo*, respectively).
- A *pronunciation model* formed by the necessary vocabulary to define the grammar phrases. The pronunciation is represented as the sequence of phonemes

corresponding to each word. The vocabulary in this game is formed by 60 direct and indirect words.

- *An acoustic model*, this acoustic model was trained with a voice corpus, designed with a linguistic content, which is based on sentences formed with phrases with linguistic meaning that contain the most of the Spanish phonemes from Mexico, this acoustic model was created and used in [7]. The acoustic model was trained with voice corpus formed by 40 people (23 men and 17 women) from different states from the Mexican Republic with mother language is the Spanish talked in Mexico [3]. The phonetic models are independent from the context and identify the phonemes from the Spanish talked in Mexico. For the training of this model were used the libraries from Hidden Markov Tool Kit libraries, HTK [8].

4 Integration between the Electronic Game and Speech Recognition System

The electronic game (memory game) was developed in the Language C and with the SDL graphic libraries. The functions of the speech recognition system was developed in Language C too, these functions contains scripts with HTK's commands. The electronic game with voice recognition runs in Linux.

The interaction between both systems start at the moment to select two stones with equal objects, in that moment the recognition of voice process starts and the player has three chances to pronounce on the correct way the object name. Words are detected with a microphone.

In each try, the function denominated *reconocePalabra* is called. The parameter that it uses is the word that corresponds to the found object name. This function does the next steps:

1. Execute the application *rec_alsa*, which saves the audio input in a format suitable to HTK. This file is called *palabra.sig*
2. Execute the application *mi_rec_htk.sh*, this function used the command HVite from HTK, the language model, the pronunciation model, the acoustic models and the file *palabra.rec* with the 5 words with most probabilities to do pronounced and their probabilities. Lately a file is generated that only has the 5 recognized words and stored them in a file denominated *palabra2.rec*
3. Execute the application *mi_rec_busca.sh*, this function receive as parameter the object found name. Its homework is to look for the parameter that receive in the file *palabra2.rec*, return 0 or 1 if the parameter was found or not in the file from the recognized words, respectively.
4. The value return *mi_rec_busca.sh* is denied and that value return to the application *reconocePalabra* to the electronic game.

Finally, if the word pronounced corresponds to the object found in the memorama game, the player takes the assigned points.

5 Results

Actually We are freeing up the final version of this electronic game with speech recognition soon, to be used for language-problem children at MAC from Tizimin, Yucatan. Once we have finished this game, we want teachers from MAC uses it from 3 to 4 months continuously to verify its usefulness in children's rehabilitation. However, we have made proofs with a beta version of the software and the results obtained are this:

- *An nice environment*, children likes the environment of this game; they get excited when they select their own characters. Children are interested in the game. This could be proved when the most 5 mischievous children, according to the their teachers, played the game. They paid a lot of attention in what was happening with the characters in the game.
- *A challenge to defeat*, children made a big effort to pronounce the words correctly; even they talk louder or repeat the words several times because they wanted to see the animation.
- *Satisfactory voice recognition*, this system was proved in children with different levels of language problems, from children with an acceptable pronunciation to children who use brackets or have a lower pronunciation than expected. In the most of the cases the system recognized correctly a lot of words pronounced.
- *A funny and addictive game*, children were excited about using the game; even when they were asked to give the opportunity to other children, they didn't want to stop playing.

This results were obtained in a qualitative way by watching 8 children (of both sex) when they interacted with the game, and in a complementary way with the comments of the 3 teachers of MAC.

6 Conclusions

In this word we described the characteristics, history and characters from an electronic game (memory game) with speech recognition, as well as the way people can interact with this game.

The speech recognition system can recognize the phonemes from Mexican Spanish, These were acoustically modelled using the context-independent modelling, and were trained using Hidden Markov Models. This continues Spanish speech recognition system is multispeaker and it uses a 60 words vocabulary (both direct and indirect words), which is used in children's rehabilitation. We presented the way in which the Memorama game was integrated with the speech-recognizing system, too.

Furthermore, it was presented the preliminary results by using the game in children with language problems. The results obtained are: *the game is funny and addictive, it has an attractive environment, play with it is a challenge to defeat and the voice recognition with children with language problems was satisfactory*. We pretend to corroborate and increase these positive results by doing in the next months quantitative and qualitative evaluations at the moment of using the game in a daily way in the MAC in Tizimin, Yucatan.

Finally, our goal is this software can be a funny, entertaining and less boring tool used to help children to repeat the words in their process of rehabilitation and evaluation. The teachers of CAM say that this tool will be very useful for children.

Acknowledgments. We thanks to CONACYT and Gobierno del Estado de Yucatan all the facilities, supports and resources given to the achiviemnt of our activities. This project is financed by Fondos Mixtos CONACYT-Gobierno del Estado de Yucatan, code YUC-2004-C03-028.

References

1. <http://www.nichcy.org/pubs/spanish/fs11stxt.htm>.
2. Miranda-Palma C. and Mayora-Ibarra O. Robotic Remote Navigation by Speech Commands with Automatic Obstacles Detection. Proceedings de IASTED International Conference on Robotics and Applications 2003. Salzburgo, Austria, Junio 2003, pp. 53-57. ISBN: 0-88986-357-1, ISSN: 1027-264X. Editorial ACTA Press..
3. Miranda-Palma C., Mayora-Ibarra O. and Urraga-Serratos E. Interface with speech recognition for the robotic manipulation. Proceedings of Workshop Human-Computer Interaction being, Advances in the Computer Science. ENC 2004, Colima, México, Septiembre 2004 pp.72-78 ISBN: 970-692-170-2.
4. Pineda, L.A., Massé, A., Meza, I., Salas, M., Schwarz, E., Uraga, E. y Villaseñor, L. The DIME project. MICAI 2002: Mexican International Conference on Artificial Intelligence. Lecture Notes in Artificial Intelligence. Vol. 2313. pp 166-175. 2002
5. Romero, L., To drive by IIMAS the computer investigation. Published on September 2 of 2002 in Gaceta UNAM. No. 3569. ISBN 0188-5138. <http://www.unam.mx/gaceta>
6. <http://www.xtec.es/~jlagares/f2kexp.htm>
7. Miranda-Palma, C., (2003), Navigation Robotic System by Means of Voice Commands with Automatic Detection of Obstacles. Mastery Degree Masteries, ITESM, Campus Cuernavaca, Computer Sciences Mastery degree.
8. Young, S. The HTK Book. Cambridge University Engineering Department, December 2002.