Using Games to Assess Oesophageal Voice

Ibon Ruiz Oleagordia, Begoña García Zapiain, Amaia Méndez
Deustotech-Life Research Unit
Avda. Universidades 24, 48007
Bilbao, Spain
E-mail: {ibruiz, mbgarciazapi, amaia.mendez}@deusto.es

ABSTRACT

This work deals with the design and development of a multimedia game for the assessment of oesophageal voice during the learning stage of the rehabilitation process. Patients who have undergone a laryngectomy as a result of larynx cancer produce oesophageal speech. The game consists of a graphic interface specially designed to use in local and remote PCs. In this application each user can access the session by introducing an identity card number and password. Therefore, each user’s database will not be changed for the data of other users. Once the user is authenticated, he/she can interact with the game. The user can record his/her own voice in different sessions. After that, the voice assessment can be obtained by measuring the most relevant parameters accepted by the international community: Pitch, Jitter, Shimmer and HNR (Harmonic to Noise Ratio).

KEYWORDS


INTRODUCTION

The second most common type of cancer is larynx cancer with a rate of 95%. Every year approximately 136,000 new cases of larynx cancer are diagnosed in the world, with an average survival rate of 5 years in 68% of the cases.

Oesophageal speech (or voice) is an alternate method for speech production without oscillation in the vocal folds. The sound is produced by releasing gases from or through the esophagus. In oesophageal speech it is thus the esophagus that oscillates in contrast to normal (laryngeal) speech where the vocal folds oscillate. Oesophageal speech is therefore speaking by eructation. Patients who have undergone a laryngectomy as a result of larynx cancer have extremely low intelligibility. This is due to the removal of their vocal fold, which forces them to use the air flowing through the esophagus.

On the one hand, some people from laryngectomized collective are not motivated to learn to speak again. On the other hand, the people who are motivated to do it, they have no tools to assess their own voice. In particular, the Biscay Laryngectomy Association (“Asociación Vizcaina de Laringectomizados”, AVL) gives some oesophageal speech courses in three levels: introductory, medium and advanced level. Nevertheless, there is no any support of Information and Communication Technologies (ICT).

One of the most important concerns for specialists in this field, and even for the patients who have undergone a laryngectomy, is the complex process necessary for rehabilitation.

Oesophageal speech is characterized by its low intelligibility, which implies that its objective measurement parameters e.g. Pitch, Jitter, Shimmer and HNR, have values outside normal ranges [1]. One of the consequences of this is the impossibility of using speech recognizers, speech to text converters or any kind of automatic response device that requires a speech signal.

Once detected the deficiency, the project would like to provide a tool the associations of laryngectomees, otorhinolaryngologists (ORL), doctors and Speech therapists in order to measure the progress made over the whole process of learning of the oesophageal voice through the Serious Games.

For this purpose, it has been developed the software for measuring the quality of oesophageal voice in the process of learning and rehabilitation. The target population is about 90,000 adults in Spain approximately. Among this collective has encountered a small number of women. The project expects to motivate to a greater number of
women who until now were isolated by its complex.

The world of games may have countless application possibilities. One of the typologies of games that have been growing in importance lately is “Serious games”, which contribute the characteristics of being social, active and value-added. Within serious games, applications like health, education, public policy and strategies can be emphasized through the functions of communication, human performance engineering or training and simulation. The authors of this paper have special interest in serious games dedicated to health. In particular, this paper describes a training game for healthcare for patient disorder treatment. In fact, although one of the main aims tends to be that of occupying leisure time and providing entertainment, in the case of the project presented in this paper, the direct application of the game is of greatest importance [2] [3] [4].

OBJECTIVES

The general objective of this work is to develop and design a game to assess the voice for people who have oesophageal speech. According to the previously mentioned points, some specific social and technological objectives appear:

Social objectives

- To improve the oesophageal voice learning process.
- To use new Information Technologies and Telecommunications (ITC) for the laryngectomized collective.
- To motivate laryngectomized women to use new technologies and to participate in laryngectomized people’s associations.

Technological objectives

- To create a database with a healthy, pathological and oesophageal voice.
- To choose acoustic objective parameters accepted by the research community, at least four parameters.
- To develop advanced digital signal processing algorithms to measure acoustic parameters.
- To validate the algorithm to any kind of voice.
- To develop an accessible user interface including developed algorithms.

METHODS

As it can be seen in all the games, the developed game has four important characteristics: display, interaction, acquisition and simulation and communication and networking. Next the technologies used are detailed to develop these characteristics.

Display

In the game developed, the results are an interface which has been programmed in Java language. This interface will be shown in the results section and is specially designed so that the shown information is as friendly as possible, accessible etc. In this way, laryngectomized people will be able to understand the results simply and clearly and to assess the voice. The main characteristics of Java are explained below.

Java is an object-oriented programming language developed by James Gosling and colleagues at Sun Microsystems in the early 1990s. The language, which was designed to be platform independent, is a derivative of C++ with a simpler syntax, a more robust runtime environment and simplified memory management.

Operating on multiple platforms in heterogeneous networks invalidates the traditional schemes of binary distribution, release, upgrade, patch, and so on. To survive in this jungle, the Java programming language must be architecture neutral, portable, and dynamically adaptable.

Interaction

This game has a very important characteristic: interactivity. The user introduces all the personal information by means of the computer keyboard. As well as this form of interaction, there is a microphone. Once the user is authenticated in the system, he/she will measure his voice. In order to do so, they will record the phoneme maintained "a". Once this process is finished the program will carry out a report of results giving the measured parameters.
Acquisition and Simulation

Voice assessment can be obtained by measuring the most relevant parameters accepted by the international community: Pitch, Jitter, Shimmer and HNR (Harmonic to Noise Ratio). The equations (1), (2), (3) and (4) provide the definition used for the voice parameters [5].

\[
\text{Pitch (Hz)} = \frac{1}{N} \sum_{i=1}^{N} f_i
\]

(1)

\[
\text{RAP} (%) = 100 \times \frac{\sum_{i=1}^{i=N/2} \left| \frac{T^{(i)} + T^{(i-1)}}{3} - T^{(i)} \right|}{N - 2}
\]

(2)

\[
\text{ShdB} = \frac{1}{N-1} \sum_{i=1}^{N-1} 20 \log \left( \frac{A^{(i+1)}}{A^{(i)}} \right)
\]

(3)

\[
\text{HNR} = \frac{r_p(0)}{r_{wp}(0)}
\]

(4)

The simulation of these results is presented in graphics which represent the evolution of the all parameters measured. The algorithms of these parameters are implemented in Octave which is the free version of Matlab. The problem with Octave is that it is oriented to digital signal processing but it is quite difficult present the results in graphics. The JOPAS API has been specifically designed for this purpose. The main characteristics of Octave and JOPAS are explained below.

Octave [6][7] is a high-level language for numerical calculation, whose syntax is compatible with Matlab, but is developed by the free software community.

What makes Octave different from other programming languages?

Octave is particularly oriented towards the scientific world. Among its main differences from other programming languages, the following stand out:

1. Native matrix operation.
2. Native operation with complex numbers.

These characteristics mean that scientific algorithms can be developed in a far shorter time than in other programming languages. Therefore, Octave is the ideal language for the development of digital signal processing algorithms, digital image processing, control systems, statistics, etc. Furthermore, there are a great many toolboxes which allow the user to avoid having to start from scratch when wishing to deal with a particular subject matter.

The jOPAS API [8], developed by the PAS group at the University of Deusto, permits the user to use the calculation power of Octave from a Java application.

There is a project called Octave-GTK which provides a number of characteristics similar to the jOPAS API, but this did not adjust suitably to the established aims.

The main difference between the projects is the language with which the Octave wrapper is implemented; in Octave-GTK, C is used, whereas in jOPAS, the language used is Java, the programming language our students are accustomed to working in.

With jOPAS the user can program applications in Java, assigning all the mathematical calculations in Octave. jOPAS is thus used as a uniting link between Java and Octave. Using jOPAS means that variables in Java can be converted into Octave variables, Octave functions can be executed and variables in Octave can be converted into Java.

Communication and Networking

The game consists of a graphic interface specially designed to use in local and remote PCs. In order to work with remote networks or PCs some technologies are necessary. For these tasks a web server and a database management are necessary. In fact, database management is also needed for local work. The following sections contain the detailed explanation of these technologies.

Web technologies

Tomcat is a Web server with support of servlets and JPS. It includes the Jasper compiler, which
compiles JPS turning them to Servlets. The Tomcat servlets motor often appears in combination with the Web server Apache. Tomcat can work as a Web server by itself. Since Tomcat was written in Java, it works in any operating system that has the Java virtual machine.

**MySQL**

MySQL, which stands for "My Structured Query Language", is a relational database management system (RDBMS) [9] which has more than 6 million installations [10]. The program runs as a server providing multi-user access to a number of databases.

The project's source code is available under terms of the GNU General Public License, as well as under a variety of proprietary agreements. MySQL is owned and sponsored by a single for-profit firm, the Swedish company MySQL AB, now a subsidiary of Sun Microsystems, which in turn was acquired by Oracle, which holds the copyright to most of the codebase.

MySQL is commonly used by free software projects which require a full-featured database management system, such as WordPress, phpBB and other software built on the LAMP software stack. It is also used in very high-scale World Wide Web products including Wikipedia, Google and Facebook.

**DESIGN**

Figure 1 shows the design of the game and particularly the graphic user interface. The figure shows the lowest level resources such as: low level algorithms, database and printer.

![User Interface](image)

**Figure 1: Design of the game**

As can be seen in Figure 1, the design of the game has hierarchic architecture. The game is designed modularly and the recommendations on software design are taken into account. Modular organization allows the independent arrangement of resources, thus facilitating their modification or extension.

The possible user actions that can be performed when using the application are explained next. The ideal scenario is one in which the user correctly introduces his national I.D. number and password (if they are not correct, an error message is displayed and the login screen appears again), and selects the language required for the application.

Then the user will be able to visualize the different graphs of the compiled voice parameters on the main menu screen, choosing among them by means of the dates. There will be a table on each side of the graph section, where the value parameters of the selected samples will be shown so as to analyze their progress.

If one wishes to record, the “Record” button must be pressed while one says a maintained “ah” phoneme. The recorder will automatically pause and the program will continue with the processing of the sample parameters, showing the new values in the graphs.

Another option is to print the graphs of the samples selected on a single sheet, identified by the national identity document, name and date of printed.

![Figure 2: Ideal scenario](image)
Finally, to return to the “Start” screen and exit the session, the “Close Session” button must be pressed.

RESULTS

Once the user enters the application, there are many possibilities. The user can record his/her voice and measure the voice parameters as can be seen in Figure 3. In order to carry out the voice assessment by measuring the most relevant parameters accepted by the international community -Pitch, Jitter, Shimmer and HNR-, the patients produce the “ah” phoneme.

As you can see in Figure 4, the user introduces the login and password to access the application. This step is necessary to maintain the privacy of the user working remotely.

If this step is taken beforehand, the assessment could be calculated. Figure 5 shows the assessment of the user’s pitch in a graph. The application has four tabs for each parameter: Pitch, Jitter, Shimmer and HNR.

The user does not need to be a specialist in voice and for the graphics to be more understandable, an explanation is included below it.

Along with the functional results, it is necessary to assess the efficiency of the patient learning how to use the tool developed. This has been possible thanks to an anonymous opinion survey -whose items can be seen in Table 1- carried out on a control group of forty patients from the Biscay Laryngectomy Association (AVL in Spanish), instructors and general users. They therefore had all the necessary knowledge and could also express a valid opinion on the tool’s use.

The opinion survey scores can be seen in Table 1. Each item has been evaluated between 1 and 5 points, according to the following scale:
The first three statements, which are related to game use, received an assessment score of over 4 out of 5. This reflects a positive evaluation, but also a certain degree of difficulty with the initial use of the tool, as can be seen in the second statement. Statements two and four are related. The fourth confirms the assertion previously mentioned. In this case, the score is 2, which means that the general opinion of the use of game is quite satisfactory.

Statement five assesses the design of the game taking to account the users’ point of view. Motivation and general satisfaction are assessed by statements six and seven. The score obtained in the last three statements is excellent since the minimum score is 4, which means “Agree” with the statements.

### Table 1 - Satisfaction survey results

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Strongly agree</td>
</tr>
<tr>
<td>4</td>
<td>Agree</td>
</tr>
<tr>
<td>3</td>
<td>Neutral</td>
</tr>
<tr>
<td>2</td>
<td>Disagree</td>
</tr>
<tr>
<td>1</td>
<td>Strongly disagree</td>
</tr>
</tbody>
</table>

For a proper implementation of the application, it is necessary to test different versions of the software at the Association of Laryngectomees of Biscay. As can be seen in the statement four of the survey, a lot of time is needed to master the software. In order to enhance the friendly use of the software this test has been made. It is compulsory to assess the advantages and disadvantages of application for the people of that association.

Due to the features mentioned above, a simplest user interface has been designed which is adapted to the needs and limitations of people with less abilities of Information and Communication Technologies (ICT). This interface consists of a single button that performs the function of recording or the user can go back to the main screen. This window displays a roadmap of how to record, telling the user what to do at all time as can be seen in Figure 6.
Finally, the recording process is performed. Figure 7 shows the results of the recorded voice without displaying any acoustical parameter. In contrast, the result obtained is explained and shown in detail. In order to be useful the software has been installed in the Biscay Laryngectomy Association (AVL) with touch screen as shown in figure 8.

After new user interface developed, the opinion of the laryngectomee people has been collected. The opinion survey scores can be seen in Table 2. Table 2 shows that there is a great improvement in several statements over the results obtained in the first survey, table 1. The application is much more intuitive, easiest to use and it is needed less time to master it. This assertion is concluded take into account the statements one, two and four of the second survey, as can be seen in table 2. Therefore, it can be deduced that the general degree of satisfaction of the application is excellent.

Table 2 - Satisfaction survey results

<table>
<thead>
<tr>
<th>ITEM</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- Is the game intuitive?</td>
<td>4.1</td>
</tr>
<tr>
<td>2- Is it easy to use?</td>
<td>3.5</td>
</tr>
<tr>
<td>3- Is the game useful to help understand the theoretical contents?</td>
<td>4.5</td>
</tr>
<tr>
<td>4- Is a lot of time needed to master the game?</td>
<td>2</td>
</tr>
<tr>
<td>5- Does the design enable going deeply into the contents of the application?</td>
<td>4.5</td>
</tr>
<tr>
<td>6- Do you find the game more motivating than conventional software?</td>
<td>4.5</td>
</tr>
<tr>
<td>7- General degree of satisfaction with the application.</td>
<td>4</td>
</tr>
</tbody>
</table>

CONCLUSION

The main conclusion obtained is that the patients have received the possibility of using an easy, usable and accessible tool to assess the disability of patients with great satisfaction. This statement is confirmed as can be seen in the opinion survey score in Table 1. Patients use games as a learning tool during the three levels of the courses they receive in speech rehabilitation.

Therefore, the objective of improving the oesophageal voice learning process has been completed, as the instructors of the process have given a positive evaluation of it.

Women, as well as men, in the association have enhanced their ability to use the new technologies as a way of helping to improve their quality of life. After implementing standard acoustic parameters for voice characterization in the game, we have realized that we can develop specific games to improve specific aspects such as hoarseness, low pitch voice or the excess air emission during conversation.

ACKNOWLEDGEMENT

The authors of this paper would like to thank the students Andrés Ignacio Gola and Bruno Martínez Irastorza for their help. It is also important to mention the support given by the Biscay Laryngectomy Association (AVL) whose members voluntarily lend his voices for this research, without their help it would not be possible to carry out this project. It must be pointed out the Education, University, and Research Basque Department support of the project.

REFERENCES

[4] Mostafa El-Said and Samah Mansour, “Serious Game Based Learning Creating a Triangle of Success: Play,


Ibon Ruiz Oleagordia was born in Bilbao (Spain) in 1975. He graduated in Physics at the Basque Country University, Bilbao, in 1999 and obtained his degree in Electronic Engineering at the same university in 2001. He has his doctoral thesis registered. Since 2000, Mr. Ruiz has worked as a lecturer and has been part of the Computer Architecture, Electronic, Automation and Communication Department at the University of Deusto. In 2002 he became part of Life research group at University of Deusto.

Begoña García Zapirain was born in San Sebastian (Spain) in 1970. She graduated in Communications Engineering, specializing in Telematics at the Basque Country University, Bilbao (Spain) in 1994. In 2003, she defended her doctoral thesis in the pathological speech digital processing field. After many years working for ZIV Company, in 1997 Dr. García joined the University of Deusto faculty as a lecturer in signal theory and electronics. She has been the head of the Telecommunication Department at the University of Deusto since 2002. In 2001 she created the e-Life research group at the same university, playing the role of lead researcher.

Amaia Méndez Zorrilla was born in Bilbao (Spain). She obtained her MSc in Communications Engineering from the University of Deusto, Bilbao, in 2001. Since 2003, Ms. Méndez has formed part of the Computer Architecture, Electronic, Automation and Communication Department at the University of Deusto, and in 2005 became a member of the Life research group at UD.