AppVox: An Application to Assist People with Speech Impairments in Their Speech Therapy Sessions

Cirano Gonçalves[1], Tânia Rocha[2(✉)], Arsénio Reis[2], and João Barroso[2]

[1] University of Trás-os-Montes and Alto Douro, Vila Real, Portugal
ciranoti@gmail.com
[2] INESC TEC, University of Trás-os-Montes e Alto Douro, Vila Real, Portugal
{trocha,ars,jbarroso}@utad.pt

Abstract. In this study an application to assist people with speech impairments in their speech therapy sessions is presented. AppVox simulates a vocalizer (audio stimulus feature) that can be used to train speech by repeating different words. In this paper, we aim at presenting the application as an assistive technology option and assess if this is a usable option for digital interaction for children with speech impairment. To assess the application we present a case study in which the participants were asked to perform tasks using the AppVox application. The results showed that this group of participants attained a good performance when interacting with the application.

Keywords: Assistive technologies · Usability · Speech and language impairments · Web application

1 Introduction

Assistive Technology (AT) is a fairly designation, used to identify the full range of resources and services that contribute to provide or augment functional skills of people with disabilities, thus promoting their independent life and inclusion [1].

This concept is also defined as "a wide broad of equipment, services, strategies, and practices, designed and used to ease the problems of disabled persons [2].

The expression "Assistive Technology" (AT) was created in 1988 as an important juridical element of the USA Public Law 100-407 and it was renewed in 1998 as the 1998 Assistive Technology Act (Public Law 105-394, S.2432) [3]. It is also an important element of the American with Disabilities Act (ADA), which establishes the rights of the disabled citizens in the USA and provides a legal base for the usage of public funds in the acquisition of support resources [4].

Speech and language disorders can interfere with the person's ability to understand, to be understood, or to express his thoughts. There are several causes, which can be present at birth or be revealed later in the infancy. They may also occur at any moment in life due to an accident or a disease [6–8].

Á. Rocha et al. (eds.), *Recent Advances in Information Systems and Technologies*,
Advances in Intelligent Systems and Computing 570, DOI 10.1007/978-3-319-56538-5_59

The specific area that aimed at expanding communication skills using AT is called Augmentative and Alternative Communication (AAC). AAC includes all forms of communication (other than oral speech) that are used to express thoughts, needs, wants, and ideas and is used by those with a wide range of speech and language impairments, including congenital impairments such as cerebral palsy, intellectual impairment and autism, and acquired conditions such as amyotrophic lateral sclerosis and Parkinson's disease. AAC can be a permanent addition to a person's communication or a temporary aid [10].

In this work we present an application, called AppVox, aiming at stimulate children with speech impairments by learning, differentiating and repeating new sounds while playing recreational exercises. It is believed that the development of speech of children with speech impairments greatly benefits from an early stimulation to speak [6]. Most times, this process is initiated with the treatment and support of an audiologist, but greatly depends of the parents support and the continuity of the exercises performed at home [7]. This fact raised our motivation to study and develop an assistive solution, in Portuguese language, to aid the speech training of this specific public.

This paper is organized in the following sections: background (in which the main concepts are presented, such as: speech disability, digital accessibly, usability; also we listed similar technologies to ours, focusing on their differences); case study (where we describe methods used, participants, experimental design, apparatus, results and discussion of the assessment of the application); and conclude with conclusions and future work.

2 Background

Bergin (1991) defines a person with speech and language impairments as having difficulties in interacting and communicating with others. In this context, a person with this type of disorder have difficulties in the ability to understand and be understood, and even to express their feelings. These impairments, affects mainly the ability to pronounce words clearly and understand spoken or written words. Also, it can also involve voice disorders, including tone, volume or quality. The causes of this disorder may apply at birth, childhood or later due to an illness or accident [8].

In several studies it is considered that information and communication technologies (ICT) can bring many benefits in the learning process [9] and therapy. The use of ICT is stated as a positive motivator to enhance performance and promotion of their users [11]. When referring to people with disabilities, the advantages of ICT use are underlined, as they reinforce the motivation for the interaction of computers used as assistive technologies [12–14].

In this sense, the concept of assistive technology translates into a set of features that make possible the autonomy of people with disabilities, in carrying out daily tasks, and an active participation in society [15, 16].

For the development of the Web application proposed in this paper was needed to research other similar technologies to understand how they work and list any problem of interaction if they had it.

Next, the technologies researched are presented.

A. Boca Feliz Application

The "Boca Feliz" application, by Smarty Ears, is a Portuguese vocalizer application, intended to help children, with oral motor disorder, to practice and workout the tongue and lips movements. To interact with this application, first we have to create a profile and insert the name of the user, then we have to insert the email of the therapist. After that, we can choose the level of difficulty. Specifically, the application has 15 exercises, demoed by a duck mascot, which can be used to help children with a verbal apraxia condition. Using the animation resource, the mascot duck asks the children to execute tongue and lips movements, such as: up and down tongue movements; right and left tongue movements; smiling; kissing; touch the nose with the tongues, etc. The cartoon draws of the duck's mouth are not very well done and it's not clear what movements the children should do. Some parents comments on the app store ware negative, stating that the app was lacking sound and wasn't helping in the children's speech development [17].

B. TalkBoard Application

The TalkBoard application, by Mark Ashley, is intended to assist communication as a visual alert board. The application can easily be customized with the user's images and symbols and it can be used by teachers, speech therapists and parents of people with autism, language and communication impairments, physical disability, such as cerebral paralysis. TalkBoard is very similar to a traditional vocalizer, while being easy to use and providing a wide range of customization options for each type of user. This application is only available for the iPAD and iPhone devices, which narrows down the group of possible users [18].

C. Talk it Application

Talk It application is a text-to-speech program launched by Microsoft, in 1997. It allows to customize the program's voice with thousands of different types of voices, Talk It aimed training children to pronounce words correctly, while they had fun to customize and test all kinds of voices. This program works in Spanish and English [19].

As augmentative and alternative communication (AAC) tools are important to aid communication for people with speech and language impairments, we next listed some AAC tools to compare vantages and disadvantages with tool that focus on training the right pronunciation of the word.

D. Plaphoons Application

Plaphoons application is a communication tool created by Fressa Project. The application allows to create communication screens and structure forms of symbols, letters and words to create messages. These messages can be viewed directly on the computer screen, to print or heard by synthesized or digitized voice. It can work directly as computer-mind communication screen, and activated using the mouse. There is also the option of using it with voice communicator, through voice synthesis features and/or recorded voices [20].

E. Avaz Australia Application

The Avaz Australia application was designed by the company Avaz to facilitate communication in children with special needs, such as Down syndrome, Angelmans syndrome, cerebral palsy and other speech disabilities. The Avaz lets you communicate

with another person through imagens, each image has a certain sound and a corresponding word. Thus a user can create messages and listen to them through a set of images. You can even customize your own image and the appropriate sound [21].

F. Verbally Application

The Verbally application was developed by the company with the same name. It is a comprehensive application, easy to use, ACC (Augmentative and Alternative Communication) for the iPad (Fig. 6). This application allows users to communicate quickly and effectively with little physical effort. To do this simply select the word you want, or write it manually, and the word that was wrote is played to the user [22].

After this research, we found that most of the applications are based on the Augmentative and Alternative Communication (CAA) technologies, only the Talk It application allows the training of the correct pronunciation of the words.

Next, a comparative table of the previously solutions is presented for a better perception of their similarities and differences (Table 1).

Table 1. Previous solutions comparative table

	Talk It	Plaphoons	Avaz Australia	Verbally
Operative system	Windows	Windows, Linus, Mac OSX	iOS	iOS
Licence	Non-Free	Free	Free	Free
Portuguese language	No	Yes	No	No
Pronunce assessment	No	No	No	No

Notice that the applications were not developed specifically for the training of the correct pronunciation of words, since they do not allow to evaluate the pronunciation of words in Portuguese automatically.

Next, it is presented our app proposal that intents to train and assess the correct pronunciation of Portuguese words.

3 Presenting AppVOX Application

The development of AppVox was based on the rational and context provided by a study, published by the Secretaria de Educação Continuada, Alfabetização, Diversidade e Inclusão in the portal Mac/Brasilia-DF, in 2007 [24]. In this study, guidelines were defined for the Specialized Education Service for children who use Augmented and Alternative Communication (AAC) on using alternative resources for effective communication.

Also, it was conducted a survey, regarding similar applications to the traditional vocalizer, as well as, exercises practiced by speech therapists with their users. After gather this information, we started the development of AppVox.

3.1 The Design

Several user interfaces were designed and experimented to get to the version here presented. An application test version was implemented, considering the ease of usage and understanding of children. Next are presented all feature of AppVox application.

The interfaces were designed to train the user's diction and improve the words pronounce.

The application uses voice recognition to evaluate the user pronounce and measures the user progression accordingly to the number of activities correctly executed.

In a proposed activity the user is invited to randomly select an image, as presented in right figure (Fig. 7).

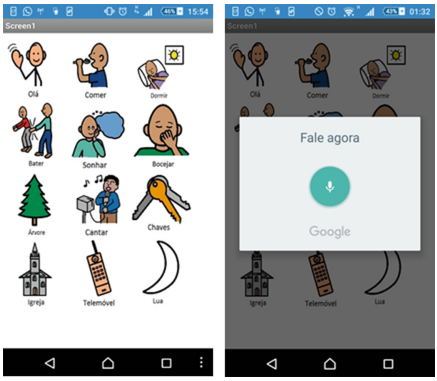

Fig. 1. AppVox interfaces.

In the application, the word associated with the image is verbally reproduced, after which, the user is invited to pronounce the word (Fig. 1). The application will wait for the user to pronounce the word and will evaluate the user performance and carry on to another activity.

This activity is meant to interact with the user by accomplishing the tasks of: clicking in an image; listen to a well pronounced word; and speak, with good pronounce, the text associated with the image.

The activity's word set is composed mainly by verbs but also includes more complex words, such as "telemovel", "chaves", "igreja" and "arvore". Thus, providing several levels of difficulty.

To create the sounds used in the exercises, we used the Google voice synthesizer. We opted for a female voice, as the female sounds are characterized by a higher pitch, and most participants can hear the high pitch sounds better than the low pitch ones.

3.2 The Implementation

The technology used to develop AppVox was the app inventor, an Android tool dedicated to developing applications and games for mobile devices that provides a

combination of programming blocks. App Inventor is an interesting tool for developing prototypes. Afterwards, the application uses algorithms written in java for android, using the ide Android eclipse, as it is a complete program, along with the use of SQLite, since this is an existing database in any installation for mobile operating system from Google and gives the developer control over large data storage.

4 AppVox Versus Other Technologies

The first application presented, BOCA FELIZ, had design features that did not help to comprehensibility of the app functions and these, for us, were disadvantages. As the duck's mouth was not very well designed, this led to a misunderstanding about the lips and mouth movements that the child should do. The parents' feedback on the application was negative, reporting that the application should include sound and also they believed that the app did not help in the speech development.

The TALK BOARD application worked only on iOS systems, is used as a voice communicator and does not have the interaction and correction of pronunciation options.

Compared to TALK BOARD, the application presented, AppVox, provides interaction between interface and the user. Besides, is an application for Android and IOS.

The TALK IT application implied entering text into an electronic device with the keyboard to interact with it before the audio stimulus. This could exclude users with severe motor impairments or limited motor control who cannot communicate by speech.

On the contrary, AppVox has a simple interface, with image buttons that allow the interaction of people with severe motor impairments. In addition, AppVox provides feedback to the user about the pronunciation of the word in order to provide interaction with the person who is using the application. Also, Talk It and Talk Board apps are paid and AppVox is for free.

PLAPHOONS application only worked in a computer, thus did not allow the user to handle it as a smartphone app, so it added the necessity to print the communication boards.

AppVox runs as a smartphone app and the second version of this app (that is being implemented) allows the user to practice as a communication board thus avoiding paper boards, feature which will bring more efficiency.

AVAZ AUSTRALIA application did not promote interaction with the user, did not encourage the user to pronounce the words and did not give feedback to the user about its performance. In addition, it was presented only for iOS. With AppVox, we offer the option of images selecting, play text through audio stimulus (synthetize voice) according to the image presented, allow users to enter voice data of the corresponding word and return feedback about to accuracy of the pronunciation.

VERBALLY application has no clear feedback option and only works on iOS.

With the AppVox, users are stimulated with audio and text feedback regarding their performance, Sentence as: "Congratulations, you hit the nail on the head" and: "Wrong word, try again" are displayed to encouraging the users to correct the pronunciation.

AppVox reinforced the concept of repetitive exercises that: combine image and sound; develop mindfulness and memory; and lead to autonomous usage of the exercises.

The AppVox application was very well accepted by the participants that interact with it, showing an increase in the group's motivation to use the application when compared to a traditional voicer or other applications.

Next, the AppVox user tests are presented.

5 Appvox User Assessment

With the user tests, we intended to assess if a group of people with speech and language impairments can use the proposed application. This preliminary assessment was made to verify if this application can be usable for this group of people, by improving their performance and satisfaction. AppVox aimed to be a natural developer of users' interactions (vocalizer and user) and an aid tool aimed to be an aid tool for speech therapy.

5.1 Participants

The participants group was composed by four children, distributed in the following categories: two children with the age of 8 years and two with the age of 10 years; two children of the masculine gender and two children of the feminine gender. Regarding academic qualifications, the 10 years old were students of the 5º grade of the 2º cycle; and the 8 years old were students of the 2º class of the 1º cycle.

The participants were selected accordingly to the average rate of literacy and primary education. Moreover, within the group, all children can read and write.

After an initial conversation with the children's parents to assess their digital integration, we confirmed that all the children had the same level of experience, regarding the user interaction with applications, i.e., using cell phones, tablets and Internet. All participants were volunteers and had permission from their parents or guardians to perform the usability tests.

5.2 Experimental Design

The participants were observed while executing task using the select action. The users had to: (1) select a button associated with an image; (2) listen to the application voicing the word associated and represented by the selected image; (3) Replay with a voice command by pronouncing the word previously listened; (4) wait for the application analysis and reply. The application compares the user voicing with a correct voicing and replies with a sound message: "Congratulations, you hit the nail on the head", if the user voicing was correct; or "Wrong word, please try again", if the user voicing wasn't good enough.

We tested this task as the user interaction with the device is of great importance, generating stimulus and motivation thought the feedback received after the user voicing.

With a traditional vocalizer it would be difficult to observe and analyze the several distinct user errors when pronouncing the words, as they did not provide feedback. For this regard, we developed the application to provide feedback to the user about his voicing performance.

The users selected the application image set, which was based on the PC program board maker with Speaking Dynamically Pro (SDP: important tool for building communication boards), and are used in the boards of the traditional vocalizers.

Regarding the evaluation criteria, we adopted the evaluation variables of usability (effectiveness, efficiency and satisfaction) to evaluate the user performance and experience. For the effectiveness, we registered how many users were able to conclude the tasks without quit. For the efficiency, we registered the resources spent to achieve effectiveness: time used to accomplish the task; errors made during the interaction with the application, in particular, the user voicing processing and recognition; and the observed difficulties, such as, the low accuracy in recognizing some words. For the satisfaction, we observed the users regarding their comfort during the tasks execution and if they accepted the task or if they asked to repeat the tasks.

5.3 Methods

In this study, the case study was allied to the usability evaluation (user tests) to assess the application.

The case study was used to assess a real application in a real context. The user tests, also known as usability tests, are used to evaluate the software interaction by allowing a group of participants to interact directly with the system [25]. As indicated above, we assess effectiveness (the capacity to accomplish the proposed task), efficiency (errors, difficulties in the interaction) and satisfaction. The methods of data collection used were directly related to the methods conducted, which includes: pre-test questionnaires, pro-test questionnaires, “think aloud”.

5.4 Procedures

The children were carefully placed in front of a mobile phone in a controlled environment and the functionalities of the application were explained to them. It was explained how to use and handle the phone and the application properly, e.g., clicking the buttons with the image to hear the message and repeat it, after which they started the activity. The order in which the children used the devices and the application was random. After explaining the task the evaluator and/or observer did not help the participant in the interaction with the application.

5.5 Apparatus

The following resources were used: a mobile phone (Sony Xperia Z1) with the Android system (was adopted due to its ease of transportation, popularity, and the potential to attract the attention of children), with a high-resolution large display of 5-inch and Full HD engine, with X-Reality for mobile devices, providing progressive analysis 1920 × 1080, for clear images, without jagged edges.

5.6 Results and Discussion

In this section we present and analyze the results regarding: effectiveness, efficiency and satisfaction.

Regarding effectiveness (execution of tasks without giving up): all participants completed successfully the tasks proposed.

Regarding efficiency (spent resources: time to complete the task, the mistakes made and difficulties encountered), we present the time spent by users in performing the tasks, with the two input devices, including the reference values:

• Of the four users, one spent more time to complete the task with the application; the faster participant spent 7 s to complete a task of selecting the image, hear the word and successfully pronounce it back; the slower spent about 15 s to successfully perform the task. There was a difference between the use of each application image button.

• Participants 8 years old experienced more difficulties in some words and to obtain a positive response in the words: yawning, tree and cellphone.

• Participants 10 years old were able to perform with minimal pronunciation errors and/or when the result was obtained by consecutively "try again". They managed to work around the problem by selecting the button and finally getting the result "Congratulations, correct word".

• During the experiment, 9 errors of 8-years-old children and 5 errors of 10-years-old children were recorded, at an average of 10 min of application usage by each one of the participants.

Regarding satisfaction (comfort and application acceptance), it was found that users liked to interact with the application, as they were smiling during the interaction and celebrating the successes. On the occasions in which they succeeded, after trying again, it was visible their joy for having completed the task. They all finished by completing all the tasks in all the buttons.

At the end of the experiment two users asked if they could repeat the task with the application. It is worth noting that the tasks were carried out randomly, but participants interacted with the device, regardless of the task order.

6 Conclusion

We conclude that the application can be used as a support tool to perform vocal exercises.

It can help speech therapist, parents or teachers in assessing which words the patient is having more difficulty to distinguish, and in the form of help that can be given

to the patient in order to improve his/her pronunciation. The application helps to encourage the child to perform the exercises, and offers parents some support for the repetition of the exercises at home.

Regarding the usability assessment, we obtain positive feedback on the efficiency, effectiveness and satisfaction in the interaction with the application however we acknowledge the need to perform more tests with more participants.

As future work we are planning to add new features to the AppVox application, such as: board and graphic symbols customization for different contexts, providing multitude of contexts, thus improving the user's cognitive development. In addition to provide feedback, the application should also generate a report with the words or phonemes with more errors, helping the pronounce treatment and correction. Also, we will develop the features that can substitute the function of vocal communication or other augmentative and alternative communication technologies. Moreover we intend to perform a complete usability assessment with these new features with more participants.

Acknowledgments. Part of this work is funded by the 2015 Digital Inclusion and Literacy Prize - "Metáfora de interação acessível para navegação Web sem recurso a texto", promoted by the Portuguese ICT and Society Network, granted by the Portuguese Foundation for Science and Technology (FCT).

References

1. Rocha T.: Accessibility and Usability on the Internet for People with Intellectual Disabilities. Master degree, University of Trás-os- Montes and Alto Douro, Vila Real, Portugal (2008)
2. Rocha, T., Bessa, M., Gonçalves, M., Cabral, L., Godinho, F., Peres, E., Reis, M., Magalhães, L., Chalmers, A.: The recognition of web pages' hyperlinks by people with intellectual disabilities: an evaluation study. J. Appl. Res. Intellect. Disabil. **25**(6), 542–552 (2012). doi:10.1111/j.1468-3148.2012.00700.x
3. Section 508. Assistive Technology Act of 1998. https://www.section508.gov/assistive-technology-act-1998
4. ADA. Americans With Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 328. https://www.ada.gov/cguide.htm
5. Araújo, A.: Jogos Computacionais Fonoarticulatórios para Crianças com Deficiência Auditiva. Ph.D. Thesis. UNICAMP (2000)
6. Oliveira, A., Almeida, E., Oliveira, S., Pinto, D'alma, A.: Como brincam as crianças surdas: um estudo luz da fonoaudiologia. In: Revista de Psicologia. Vetor Editora, vol. 7, no. 2, pp. 77–84 (2006)
7. Prado, A.C.: Principais Características da Produção Vocal do Deficiente Auditivo. In: Revista Brasileira de Otorrinolaringologia. Rev. CEFAC, São Paulo, no. 3, pp. 404–410 (2007)
8. Bergin, S.: Communication disorders: breaking through the barriers. In: Ensign, pp. 46–50 (1991)
9. Rocha, T., Bessa, M., Cabral, L.: Performing universal tasks using a mini iPad: usability assessment per people with intellectual disabilities. In: Proceedings of the Interaccion 2016 - XVII International Conference on Human Computer Interaction, Salamanca, Spain, 13–16 September (2016)

10. America Speech-Language-Hearing Association. http://www.asha.org/public/speech/disorders/AAC/
11. International Institute for Communication and development - IICD. ICTs for Education: Impact and Lessons Learned from IICD - Supported Activities (The Hague: IICD, 2007). http://www.iicd.org/files/icts-for-education.pdf
12. Gutterman, B., Rahman, S., Supelano, J., Thies, L., Yang, M.: Information Communication & Technology (ICT) in Education for Development (2009). http://unpan1.un.org/intradoc/groups/public/documents/gaid/unpan034975.pdf
13. Kirinić, V., Vidaček-Hainš V., Kovačić, A.: Computers in education of children with intellectual and related developmental disorders. In: Proceedings of Computers and Education Conference of 32nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2009, Opatija, Croatia, 25–29 May, pp. 39–43 (2009)
14. European Commission: Educação em tecnologias de Apoio para Utilizadores Finais - Linhas de Orientação para Formadores, Comissão Europeia DG XIII, Programa de Aplicações Telemáticas - Sector Deficientes e Idosos. DE3402/Eustat project, Deliverable D06.3 (1999)
15. British Educational Communications and Technology Agency - BECTA.: The impact of ICT in Schools: Landscape Review. Becta, Coventry (2007). http://dera.ioe.ac.uk/1627/
16. Ribeiro, J., Almeida A.M., Moreira, A.: A utilização das TIC na Educação de Alunos com Necessidades Educativas Especiais: resultados da aplicação piloto do inquérito nacional a Coordenadores TIC/PTE. Indagatio Didactica, vol. 1, nº. 2 (2010). ISSN: 1647-3582
17. Boca Feliz. http://www.pucrio.br/pibic/relatorio_resumo2014/relatorios_pdf/ctch/ART/ARTThiago%20Macedo%20e%20Camila%20Rabelo.pdf
18. Talk board. http://cintiappires.blogspot.com.br/2012/09/comunicacao-alternativa-utilizando.html
19. Talkit. http://talkit.software.informer.com
20. Tecno Accesible. www.tecnoaccesible.net/content/plaphoons
21. Avaz. http://www.avazapp.com/
22. Verbally. https://itunes.apple.com/pt/app/verbally/id418671377?mt=8
23. Rocha, T.: Interaction metaphor for Access to Digital Information an autonomous form for People with Intellectual Disabilities. Ph.D. Thesis, University of Trás-os-Montes and Alto Douro Vila Real (2014)
24. Secretaria de Educação Continuada, Alfabetização, Diversidade e Inclusão. In: The portal Mac/Brasilia-DF, pp. 57–80 (2007). http://portal.mec.gov.br/seesp/arquivos/pdf/aee_df.pdf
25. Fagan, J.C., Mandernach, M.A., Nelson, C.S., Paulo, J.R., Saunders, G.: Usability test results for a discovery tool in an academic library. Information Technology and Libraries, pp. 83–102 (2012). https://ejournals.bc.edu/ojs/index.php/ital/article/viewFile/1855/1745