

Application Focused on English Language Teaching for Children, with Speech Recognition and Synthesizing Capabilities

Davi Samuel Maia Dias¹, Ricardo Silva Parente², Vanessa Levy³, Antônio Estanislau Sanches⁴, Jorge de Almeida Brito Júnior⁵, Manoel Henrique Reis Nascimento⁵

^{1,2,3}Department of Technology, University Paulista - UNIP, Manaus, AM, Brazil

⁴University of the State of Amazonas (UEA), Manaus, Amazonas, Brazil

^{5,6}Research department, Institute of Technology and Education Galileo of Amazon - ITEGAM, Manaus, AM, Brazil

Abstract— This project will present an application focused on the teaching of the English language to the children, this application being an important teaching tool, where the child can begin a cycle of learning a new language, something that will be very important in their training academic, and will serve in your professional future. In addition to showing how software is being developed and the resources used in it, this project is also concerned with presenting concepts such as: foreign language learning for children, voice recognition and synthesis, intelligent systems capable of recognizing and synthesizing the voice and the Java API Speech. To aid in English studies, the application makes use of illustrative images, themes, interactive questions, training mode, speech recognition and synthesis, which contributes to the development of writing and pronunciation in the language, mainly for making use of the resources of voice, which are the strongest point of this tool.

Keywords— English; Learning; Educational; Voice and synthesis.

I. INTRODUCTION

The importance of learning a language beyond the mother tongue is one of the characteristics of the process of advancement and globalization of humanity, where the media and all kinds of technology have undergone drastic changes over time, and the labor market has also accompanied this evolution. So that more and more professionals with a higher degree of qualification are required.

According to Pati (2017) in the 53rd edition of the salary survey of Catho, where 13 thousand people were interviewed, knowing how to speak English guarantees a salary jump of up to 61% depending on the employment area, which proves the importance of this language and others in many sectors that need this type of specialty.

Learning a new language is something that requires a certain amount of time and dedication, so it is advisable to learn from an early age, especially in the infantile phase, so that when you reach adulthood, there is no worry of not speaking a second language. According to Duarte and Batista (2013, pp. 293-301), children have a high degree of assimilation, they can absorb content quickly and practically, and they usually have more time available than many adults. the best phase to start learning, including a new language.

Knowing one of the most talked about and important languages in the world, as mentioned above, is extremely significant in the current scenario, but many students do not value this kind of study because English is not the official language of Brazil, even if it is present in the curriculum of many primary and secondary schools, preferring to give more importance to other fields of knowledge, which in a way will also be very useful also in their academic formation, however it is a fact that being able to speak English is requisite for many high-paying jobs, and can also guarantee many academic and exchange opportunities.

The purpose of this project is to offer an alternative that will help in the study of English, and for this reason the team that started this work started the development of a tool that aims to teach English to beginners, especially to children, offering a first contact to the user. will serve as a gateway to more complete learning of that language. The application has already proved to be promising, making use of synthesis and voice recognition, which is its main resource so far, in addition to others, thus helping the student with the pronunciation of words in English. However, even if the software already shows good results, the developer team feels that there is room for further improvements and implementations that can be added later by the developer group of this tool.

II. JUSTIFICATIVE

The project in question was developed due to the lack of software of this type and in order to help and contribute to the basic teaching of the language, so that people who do not speak English (especially children) have a first contact with the language in a fun way and productive, as well as giving an incentive in the study of English of the people who use this tool.

The prototype of the application presented here serves as a first contact with this language, which happens in a relaxed way, facilitating and giving an additional stimulus to the user of the tool, so that it can enter the sphere of knowledge of the English language, since English is a universal and fundamental language for people today, so this first contact is of the utmost importance because it must be something cool, fun and interesting to the beginner in English. Combining all this with speech synthesis technology and speech recognition technologies that are current technologies that facilitate the learning of pronunciation, it makes the project more adequate to what this team of students aims for, thus contributing to the foreign language teaching system in a way effective in the schools or in the pupil's own house, as a kind of aid in his studies.

III. METHODOLOGICAL PROCEDURE

The tool is being produced in the Java language with the help of Netbeans IDE 8.2, until then the Java Speech library for speech synthesis and speech recognition was used. The project whose name was adopted by the team was "SpeakApp" makes use of many colorful figures, which is a way to make the software more attractive to children.

The procedures to achieve this tool were based on the applied study of technologies such as synthesizing and voice recognition, where the knowledge obtained was applied beautifully in the system, from there it was necessary a basic analysis research on how to catch the attention of children. Another important point is to conduct tests with children, which was successful, since the forms used by the team to attract children's attention worked.

SpeakApp works with writing and pronunciation of words in English, always relating them with images, to facilitate learning. At the time this project was written, the principle to be explored, is to work with only four themes, which can be expanded in the future, so that themes are addressed: numbers, letters, animals and colors.

The tool created for this work plan aims to dynamically draw the attention of boys and girls to learn English. For this task were used good coloring drawings and a simple aspect of design, compacting with the interactivity and

the pleasure of the user to enjoy this application for the knowledge of the English language.

IV. TOOL DESIGN



Fig.1: Tool design.

Source: Authors.

In the initial screen that is the menu, are presented four representations containing each one, a theme, whose each one of the subjects can be identified by the characteristics of the image, that is well illustrative and of easy identification, besides being possible to be distinguished by the name that is above the figure. When the drawing with the title "Colors" (example) is clicked, a new screen is opened (this is the case for all themes), which will be shown below and explained accordingly.



Fig 2: Training Screen.

Source: Authors.

It is worth mentioning that there is a menu bar that contains a menu called "Options", in it is an item with the name "Students" that when clicked will show a message with some information about the components of the team that built the application and this project complete. On the right side there are the "Levels" where the user can choose the way in which he wants to start the software,

having "Alternative Mode", "Writer Mode" and "Speech Mode".

Were used very illustrative forms that draw the attention of the boys and girls, in order to make them take an interest in the software already in the menu screen, even before the use of fact of the tool.

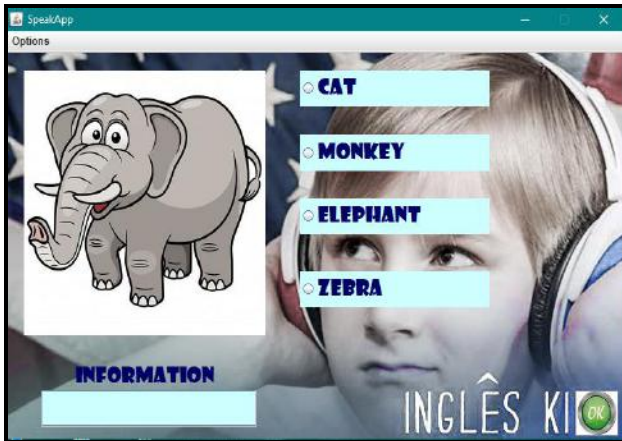


Fig 3: Alternative Mode.

Source: Authors.

V. LEARNING OF FOREIGN LANGUAGES FOR CHILDREN

Researchers in the field of neuroscience have indicated that the ideal age for language learning occurs in the first ten years of life, according to theorists such as Penfield and Roberts (1967, DIMER, SOARES, 2012, p.53). In this stage of life the brain is able to present a high degree of plasticity, this period being the highest point of this peak, and in puberty the brain no longer reaches these same capacities, because they are gradually lost.

According to Castro (1996), it was once believed that initiating a second language at the stage of literacy might be detrimental to the development of the mother tongue.

"The cerebral availability obtained in childhood, according to some studies, will never be obtained again. In addition, up to ten years of age, the number of synapses (neural connections) in the human brain remains stable (increasing gradually), as adolescence, the proportion of synapses is reversed, which also suggests less facility for acquiring language after the first ten years of life" (DIMER; SOARES, 2012, page 53).

Children have a remarkable greater ease of learning, and therefore tend to show greater progress in pronunciation, comprehension and storytelling. Children exposed to a foreign language acquire fluency faster than an adult because they have greater phonological control than older individuals. (DIMER, SOARES, 2012). "At 12 months of age, babies have a vocabulary of up to 50 words, but by the age of six it can reach about 5,000 words" (BRIGGS, 2013).

In the teaching of a language one must take into account the age issue, since children, adolescents and adults have different learning characteristics, and because of this fact, different methods of approaches must be made for each age group, always seeking the best suitability for the study, in order for the student to be able to adapt to the language taught (LIMA, 2008, pp. 297-298).

VI. VOICE RECOGNITION

Speech recognition is a set of techniques with the objective of transforming oral language into a written text, so that with this text the computer or apparatus through software, can perform some desired task using the data obtained by voice recognition.

For an application to effectively do voice recognition, it digitizes speech through a mechanism, converting the vibrations provoked by speech into digital data, this is a kind of analog-to-digital conversion. To avoid noise in the audio, the scanned sound needs to be filtered, thus leaving only the part of the sound that matters, thus eliminating external noise and interference (PEREIRA, 2009).

Then the computation of the frequency characteristics of the voice (spectral domain) is performed, so that it can be synchronized to its classification, where the sound digitization needs to separate the audio into small phonetic parts of the size of a syllable, so that the comparison with a database can be made, and thus identify what is said in the small fractions of sound. In the end, the parts are joined together forming words (PEREIRA, 2009).

Recognizing speech is an alternative to typing, this offers many benefits to the user, from the convenience of registering a text without having to type until the verification of the pronunciation of a sentence in another language, which helps in learning a new language, and many people with physical and visual disabilities, unable to type something into a computer, can make use of and benefit from this type of technology (WHAT IS SOFTWARE ..., 2018).

VII. VOICE SYNTHESIZATION

Speech synthesis is the conversion of written text into spoken language. Speech synthesis can also be referenced as the TTS (text-to-speech) conversation. Because the speech is being produced through an electronic device, it is an artificial voice that imitates human speech (MARANGONI; PRECIPITO, 2006, page 5-6).

Computers work basically in three stages (input, processing and output), voice synthesis is a form of output, the computer or any other electronic device that makes use of it, uses features such as loudspeakers to offer this kind of output (SUMMARY ..., 2018).

This way you can achieve a multitude of desired types of results for various types of tasks that benefit from this feature, such as learning the pronunciation of words in a new language or helping people with visual impairment to listen to what the computer says, are possible with the aid of speech synthesis.

In order for the computer to be able to synthesize voice some steps must be followed, among them are: Analysis of text structure, text preprocessing, text to phoneme conversion, prosody analysis and waveform production. Within these stages paragraphs, sentences, punctuations, abbreviations, acronyms, dates, times and numbers must be analyzed so that the phonemes are generated for each word of the text, and thus produce a speech with correct rhythm and intonation for each textual occasion (MARANGONI; PRECIPITUS, 2006, pp. 5-6).

VIII. INTELLIGENT SYSTEMS ABLE TO RECOGNIZE AND VOICE SYNTHESIZE

According to Monteiro (2010), recognizing and understanding speech is something that human beings have been developing since the earliest times, human speech is an intelligent means of communication that enabled the evolution of them, being humans considered intelligent beings by this and for other reasons. Over time new techniques and forms of modern communications have been made, to the point where machines with the aid of software have also begun to recognize and even understand the language spoken by man, increasingly passes to be with. Nowadays it is possible to find intelligent personal voice assistants such as Siri (Apple), Cortana (Microsoft), Google Now (Google / Android) and S Voice (Samsung) (STANDARD, 2016).

Through processing after the capture of a natural language, it is possible for the computer to recognize words and even voice commands, as mentioned earlier, being a technique used by some intelligent systems, which somehow recognize the speech pattern. There are three levels of speech recognition (recognizes natural speech), discrete (recognizes spoken speech and pauses between words) and commands (recognizes a very large number of words) (STAIRS; REYNOLDS, 2006 apud GOMES, 2010, page 243).

IX. JAVA SPEECH

In the present application, the Java Speech API is used, which is a tool created to enable speech recognition and synthesis of Java applications.

Sun has defined specifications that represent a generic interface to an engine, the Java Speech API (JSAPI). JSAPI works as a layer between programs and engines that are developed by third parties. The engines are very important because they work with the sound card by

capturing the audio (speech) or synthesizing a text (CASTILHO, 2008).

X. SCHEDULE

Table.1: Schedule

Description of Steps	AUG	SEPT	OCT	NOV	DEC
Literature review	X	X	----	----	----
Data collect	X	X	----	----	---
Data analysis and synthesis elaboration	----	X	X	X	---
First writing and correction	----	X	X	----	---
Delivery of the project	----	----	----	X	---

Source: Authors.

XI. CONCLUSION

The tool performed well and achieved great results, the satisfaction of those who used it was positive. The application is modular and proposes to be interactive in order to involve the child in the learning of the English language, collaborating to the maximum for the ease of handling and help of the teacher.

The software has a good synthesis and recognizes the speech and pronunciation of the user, thus obtaining acceptance of the use of the tool as a learning aid.

ACKNOWLEDGEMENTS

The Paulista University (UNIP). To the Institute of Technology and Education Galileo of the Amazon (ITEGAM), to the University of the State of Amazonas (UEA) Manaus/Amazonas

REFERENCES

- [1] BRIGGS, H. Scientists discover why children can easily learn more than one language. BBC News - Brazil, 2013. Available at: <https://www.bbc.com/portuguese/noticias/2013/10/131009_linguagem_infancia_an>. Access date: September 5, 2018 Perfect, T. J., & Schwartz, B. L. (Eds.) (2002). Applied metacognition Retrieved from <http://www.questia.com/read/107598848>
- [2] CASTILHO, M. P. Article Java Magazine 04 - JavaSpeech. DevMedia, 2008. Available at: <<https://www.devmedia.com.br/artigo-java-magazine-04-javaspeech/8916>>. Access date: September 11, 2018.
- [3] DIMER, D. L. ; SOARES, A. English language teaching for children. Revista EnsiQIopédia -

- FACOS / CNEC Osório Vol.9 - Nº 1 - October 2012 - ISSN 1984-9125. Available at: <http://facos.edu.br/publicacoes/revistas/ensiqlopedia/outubro_2012/pdf/o_ensino_de_lingua_inglesa_para_crianças.pdf>. Access date: September 5, 2018.
- [4] DUARTE, B. S .; BATISTA, C. V. M. Child Development: Importance of Operational Activities in Early Childhood Education. XVI Education Week. State University of Londrina - UEL, 2013. Available at: <www.uel.br/eventos/semanaeducacao/pages/arquivos/ANAIS/ARTIGO/SABERES%20E%20PRACTICAS/DESENVOLVIMENTO%20INFANTIL.pdf>. Accessed: September 3, 2018.
- [5] GOOGLE Voice Synthesis, how does it work ?. Cell Phones, 2018. Available at: <<https://www.telefonescelulares.com.br/sintese-voz-google-como-funciona/>>. Access date: September 9, 2018.
- [6] GOMES, D. S. Artificial Intelligence: Concepts and Applications. Revista Olhar Científico - Faculdades Associadas de Ariquemes - V. 01, n.2, Aug./Dec. 2010. Available at: <<http://www.olharcientifico.kinghost.net/index.php/olhar/article/view/49/37>>. Access date: September 10, 2018.
- [7] LIMA, Ana. Foreign language teaching for children: the role of the teacher. Cadernos da Pedagogia - Year 2, Vol.2, No.3 jan./jul 2008. Available in: <<http://www.cadernosdapedagogia.ufscar.br/index.php/cp/article/view/48/41>> . Access date: September 5, 2018.
- [8] MARANGONI, J. B .; PRECIPITO, W. B. Speech Recognition and Synthesizing Using Java Speech. Scientific and Electronic Journals - FAEF, 2006. Available at: <http://www.faeef.revista.inf.br/imagens_arquivos/arquivos_destaque/bjMnA2Zwc9685z8_2013-5-27-15-40-25.pdf>. Access date: September 8, 2018.
- [9] MONTEIRO, J. Evolution of Communication in Organizations. Administradores, 2010. Available at: <<http://www.administradores.com.br/artigos/economia-e-financas/evolucao-da-comunicacao-nas-organizacoes/43279/>>. Access date: September 10, 2018.
- [10] WHAT SOFTWARE IS A SOUND RECORDER AND WHAT DOES IT WORK ?. Just ask gemalto, 2018. Available at: <<https://www.justaskgemalto.com/us/en-us-and-software-of-recognition-of-voice-and-as-functions/>>. Access date: September 7, 2018.
- [11] STANDARD, M. Voice assistant is the future, but it needs to get even smarter. Uol - Notices Technologies, 2016. Available at: <<https://tecnologia.uol.com.br/listas/reconhecimento-de-voz-ganha-forca-na-tecnologia-oque-falta-para-melhorar.htm>>. Access date: September 10, 2018.
- [12] PATI, C. Speaking English increases your salary and this research shows how much. Exame, 2017. Available at: <<https://exame.abril.com.br/carreira/falar-ingles-aumenta-seu-salario-e-esta-pesquisa-mostra-ouquanto/>>. Access date: September 2, 2018.
- [13] PEREIRA, A. P. How does speech recognition work ?. TecMundo, 2009. Available in: <<https://www.tecmundo.com.br/curiosidade/3144-como-funcion-o-reconhecimento-de-voz-.htm>>. Access date: September 7, 2018.