

Software application to convert Basque texts from the Web to speech

Imanol Madariaga, Veronica Nieves

University of the Basque Country

imanol@bips.bi.ehu.es

Abstract

This work describes a tool designed to read Basque texts on the Web via a TTS system. Implemented as a button, it can read any Basque text displayed by the web browser.

Keywords: text-to-speech converter, web browser, toolbar

1. Introduction

Since the invention of the computer back in the 20th century, the human being has endlessly tried to develop that tool. In this environment society's challenge for the last years has been to provide those machines with the human capabilities to read, write or speak.

The efforts made during the last decades in this field have led in nowadays' speech technologies after a long scientific and technological development process. Within these technologies based on oral language, there are four main lines of research (OESI, 2008):

- Speech recognition: the objective of this technology is to find a comfortable speech-based interface for a perfect human-computer interaction.
- Speech synthesis or text-to-speech conversion: speech messages not previously recorded can be generated by means of this process. Introducing a text, these systems automatically convert it producing a voice as humanly and natural as possible.
- Speech generation or dialogue management: speech replies are generated with this technique using templates or sometimes signals recorded beforehand.
- Speaker verification: our voices are distinctive, and thus this technology is used to automatically verify by means of the voice the identity of the speaker.

Many institutions work in this field in order to meet the needs of modern society.

This project fits within a research line carried out by the Aholab Group at the University of the Basque Country.

In the following pages, a description of a tool to make the use of the AhoTTS text-to-speech converter by the Aholab Group easier follows: it is a software application to fit within an Internet browser that provides capability to read any Basque text on the Web.

2. System design

The main objective of the developed application is to carry out the synthesis of texts on the Web written in Basque and to reproduce the selected signals.

The tool needs to deliver the best performance from the moment the application starts to run, accepting only the quickest time to reproduce the sound. This is why design needs to be as simple as possible, using little resources.

2.1. AhoTTS

The AhoTTS text-to-speech converter developed by the Aholab Group will be used to synthesize artificial speech. The group has conducted deep research in the field of speech technologies, providing users with several tools. For instance, the on-line AhoTTS demo (Aholab, 2006) used for this work.

2.2. Internal design

As mentioned above, the tool is intended to provide Internet users with a service. Therefore, its development environment will be the Internet browsers (Internet Explorer, Mozilla Firefox, Opera...), whichever system operator is used. Thus the application is based on HTML and JavaScript languages.

The first step is to select a text to be read by the application. The text will be sent to the server via functions accepted by the browser, and there the speech is generated. If no text is selected, the application has been programmed to display the window in Figure 1:

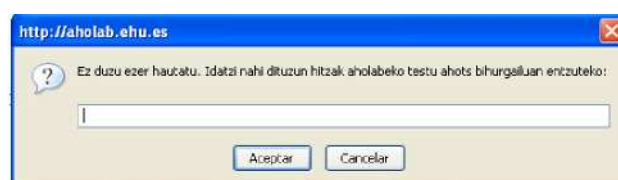


Figure 1: Warning window

The user will then know that the text has not been read, and it can be typed in.

The AhoTTS converter contains numerous parameters to control the configuration of the sound, the signal and the reproduction. For the interface to be as simple as possible, the parameters have been previously configured in order to generate quality sound without requiring anything from the user. In that way the use of the tool is made easier for non-experts in technologies. The GET HTTP method is used to send those parameters and the text to be synthesized to the server. After generating the speech, the server program sends it to the user's browser.

The Aholab text-to-speech converter generates the wav audio signal at the server. In order for the sending from the Web to the user's computer to be quicker the audio file is compressed to MP3 format with the LAME independent software. Even if the file is then ten times smaller, the amount of time needed to download a complete audio file when listening to long texts can be very long. Streaming technique is applied to reduce the uncomfortable waiting time, that is, we will start to hear the text without waiting for the complete audio file to reach. For that purpose, the XSPF Web Music Player independent software program sends the file via Flash technology from the server, and the Flash player manages it in the user's computer. It is necessary to have the Flash player installed in the user's browser (generally it is installed, and the Adobe Enterprise offers Flash Player for free).

2.3. External design

2.3.1. The button

The application is built to be located in the user's browser, in an easy-to-use icon that occupies little space. The chosen implementation is a button labelled "AhoTTS". The button has to be clicked after selecting the text in order to listen to the speech.

Figure 2 below shows a captured image of the browser and the location of the button:



Figure 2: The AhoTTS button in the browser

During the whole design process, it was clear that different browsers had to be tested. Some parts of the code behaved differently depending on the browser used. That is why, even if the application has been developed on the same basis for all versions, some small changes in the code were necessary to be able to work with different browsers. Hence two versions of the button have been developed, one to work with the different versions of the Internet Explorer browser, and the other for Mozilla, Firefox and Opera.

2.3.2. The web player window

Clicking the AhoTTS button, the XSPF Web Music Player opens in a small separate window. For its graphic design, we considered:

- Fixed measures: 300 pixels high and wide. The user cannot change the window measures.
- The browser tool: for an easy usage, the tools integrated in the browser are not used: neither the scrollbar, nor the address bar nor any other bar. The only exception is the title of the browser window: "AhoTTS".
- Elements: there are three elements inside the window: a short sentence as the entry point, the Aholab logo and the web player.



Figure 3: The player window

The following features have been considered for the web player:

- Fixed measures: 15 pixels high and 260 pixels wide. The user cannot change the window measures.
- Moving sentence: the sentence "AhoTTS, Aholab laborategiko testu ahots bihurtailua (AhoTTS, text-to-speech converter from the Aholab laboratory)" appears in motion.
- Automaticity: the audio file is automatically heard when the window opens.
- Basic functions: it is possible to stop or start the reproduction of the file, and to control the sound volume.

All that is shown in Figure 3.

3. User's handbook

3.1. Installation

The first step for the installation is to select a suitable version from the Aholab website¹. Then make sure that the Links toolbar is visible. Drag the AhoTTS link to the toolbar and leave it there. The AhoTTS button will appear in your browser, as shown in Figure 2.

3.2. Uninstallation

Right click on the AhoTTS button and use "Delete ('Eliminar')" or "Remove this application... ('Quitar de la barra')" to eliminate the application from the toolbar.

Figure 4 shows an example in Firefox.

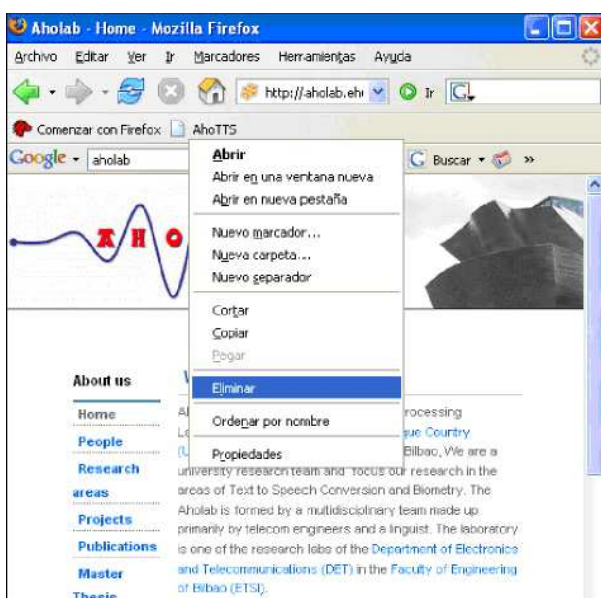


Figure 4: To uninstall the AhoTTS button

3.3. Usage

3.3.1. Texts on the Web

The application is designed to be used with a mouse. Select a text from the Web and click the AhoTTS button to listen to it. If there are no errors, a new window opens for the player. There the sound corresponding to the selected text is automatically heard.

It is possible to use these three functions:

- Stop (⏏): to stop the reproduction of the audio file.

- Play (▶): to reproduce the audio file from the start.
- Volume configuration (🔊): click the icon to change the intensity of the sound.

3.3.2. Keyboarded texts

In order to listen to the sound corresponding to a text that is not on the Web, click AhoTTS without selecting a text. The warning window shown in Figure 1 opens, and there the desired text can be typed in.

If the button is clicked by error without having selected a text, the window can be closed clicking [X].

4. Conclusions

This paper makes the use of the AhoTTS converter easier and describes a tool that serves to reproduce any Basque text on the Web in a suitable and simple way. The application is designed to be integrated in a web browser and should be useful for many users and purposes; for instance, for visually impaired people or to read web pages with no effort for the eyes, to learn how to pronounce Basque, to listen to a text while looking at something different on the screen, or to let others listen to a text when we cannot read it ourselves.

5. References

- Aholab website (2002). http://aholab.ehu.es/ahoweb/eu_index.html
- OESI, Oficina española en la sociedad de la información, tecnologías lingüísticas (2008). <http://oesi.cervantes.es/oesi/tls.jsp>
- Lemmetty, S. (1999). *Review of Speech Synthesis Technology*. Helsinki: University of Technology.
- Pardo, J. M. *Nuevas fronteras de la tecnología del habla: síntesis de voz con emociones*. Madrid: Polytechnic University.
- Cienciadigital website <http://www.cienciadigital.es/hemeroteca/reportaje.php?id=66>
- XSPF Web Music Player (2006). <http://musicplayer.sourceforge.net>
- AhoTTS Web Demo (2006). http://aholab.ehu.es/tts/tts_en.html
- Gundavaram, S. (1996). *CGI Programming on the World Wide Web*. Sebastopol, California: O'Reilly & Associates, Inc.
- Gurlekian, J. A.; Torres, H.; Colantoni, L. (2004). Modelos de entonación analítico y fonético-fonológico aplicados a una base de datos del español de Buenos Aires. *Estudios de Fonética Experimental* 13: 275-302.

¹ <http://aholab.ehu.es/aholab/Otras-demos/AhoTTS-in-your-browser.html>