

Changes to Spanish Festival synthesizer with adaptive automaton

Responsible Researcher: Claudia Zapata

Coresearchers: Rosalía Caya and César Ruiz

Department of Engineering -Informatics Engineering Section

The production of an automatic voice through a computer is a huge interesting field for many research centers, not only for the huge impact caused to achieve a very similar approach to human speaking, but also for implying a comprehensive cooperation work of several specialties.

Festival software was modified twice to provide, to the Spanish language contained therein, characteristics belonging to the way of speaking of our language: pauses, melody patterns and appropriate number reading (time, cardinal numbers, coins and dates).

Method used

To design modifications to *FESTIVAL* synthesizer, adaptive automaton uses a sequence of states and reactions for each specific condition. In this case, the condition for entry is a correctly written text which must be reproduced, and the sequence of states is how this text must be reproduced orally. (See picture 2)

Improvements made

- Pauses due to punctuation marks: The duration of the pause is determined according to the punctuation mark used.
- Pauses due to language use: Upon reading a sentence, no pause is made between each written word but a sequence of words is pronounced before pausing. This sequence is known as a phonologic word. For example: the phrase *El tigre se subió en el árbol* [The tiger climbed onto the tree] has 6 words and it is read as if they were only three words “Eltigre se subió en el árbol.” (See picture 1)
- Melody patterns: Intonation at the beginning or at the end of the phrases depends on the fact if a certain punctuation mark is placed before or after. For example, if a sentence is between parenthesis, it should be read with low intonation in the first and last syllable: *Hacía calor* [It was hot] ↓(summer arrived very soon said year)↓ y *las ventanas estaban abiertas* [and windows were opened].
- Date reading: There are different date formats. The software was adjusted for a correct reading.

- Arabic number reading: The value of a digit varies according to the position it has within the number by multiplying it for the base raised to the position. Thus, the first digit beginning from the right has the value representing its symbol multiplied by $10^0 (=1)$. The digit placed to the left has the value representing its symbol multiplied by $10^1 (=10)$, and so on.
- Coin reading: According to the symbol before the amount, it is possible to infer which is the monetary system to which it belongs and therefore the correct manner to read it.

Results of understanding tests

A standard test was used to measure voice quality in communication devices. A test was performed to 75 people and a significant improvement in the modified software was seen.