ALBAYZIN-08 System Evaluation Plan, Text to Speech (ALBAYZIN-08 TTS)

Aim

The main goal of this evaluation is to promote the exchange of ideas, to stimulate creativity, to favor the collaboration between research teams that focus their research on TTS and to compare different techniques employed. With this purpose in mind, we propose a contest among speech synthesis systems, similar to the Blizzard Challenge for English and Mandarin. The voice development would be carried out on a common database for Spanish.

General conditions

The participating teams will have up to seven weeks to build their system from the development material that will be supplied. They will also be allowed to use any other material or subsystem (acoustic models for segmentation, extern prosodic modules, lexicon, etc) which will have to be pointed out in the description of the system, including the appropiate references. Next, a set of sentences will be released and the teams will have one week to synthesize then and send the audio files back. No manual intervention is allowed during synthesis (Prompt sculpting, Using different subsets of the database for different test sentences or sentence types unless this is a fully automatic part of your system, etc).

Each participant is expected to provide five speech experts and Spanish native speakers as listeners of the evaluation.

Evaluation corpus

The information related to the test sentences will no be public until the evaluation campaign is launched, in order to avoid any tailoring of the TTS system to them. In any case, the corpus will consist of several hundreds sentences, of which only a part will be used during the subjective evaluation in order to determine both the naturalness and the intellibility of the synthetic voices.

Development material

The corpus for development will consist of approximately an hour of recording by a female voice talent called Marta. Together with the voice waveforms, automatic segmentation labels will be supplied in sampa format. The participating groups will be free

to use directly the supplied segmentation or to generate a new one by its own means. The chosen alternative must be indicated in the description of the system.

Provisional schedule

☐ May 5th , 2008

– Publication of the detailed evaluation plan in the JTH2008 web

– Beginning of the registration term

– Registration form available in the JTH2008 web.

☐ June 15th ,2008

– Development material released

☐ July15th, 2008

– Deadline for registration

☐ September 5th, 2008

– Test sentences released

☐ September 12th, 2008 (till 24:00 GMT+2)

– Deadline for submitting synthesized speech, following the established method and format in the detailed evaluation plan

☐ September 17th, 2008

– Evaluation campaign begins

☐ October 12th, 2008

– End of Evaluation. Statistical analysis begins

☐ October 31st, 2008

– Notification of the evaluation results to each of the participants

☐ November 12-14th, 2008

– Presentation of results of the ALBAYZIN-08 TTS evaluation in the V Workshop on Speech Technology, in Bilbao

Registration

The research teams willing to participate in the evaluation campaign will have to fill in the form available in JTH2008 web since May 5th 2008. The registration deadline is the

June 15th 2008. Before registering, it is recommended to read carefully the detailed evaluation plan that will be available in the JTH2008 web since May 5th. This plan might suffer modifications, due to restrictions, needs not foreseen initially or errors, which in any case will be communicated to all the participating groups. The interested ones can go directly to the organizing committee to consult aspects that are not clear in the evaluation, by e-mail to eva.navas@ehu.es and by phone (946017306).

Commitments of participating teams

1. To use the data provided exclusively with research aims. Participants will be able to use the database to develop or evaluate their own systems, but they will not be allowed to distribute nor sell them. They will have to reference the database as follows:

    Ignacio Hernández, Asunción Moreno. Diseño de un corpus para una base de síntesis de voz. Actas del XV Symposium Nacional de la U.R.S.I., Zaragoza (España), Septiembre 2000.

2. Each participant will be expected to submit in addition to the synthesized waves, a paper (preferably in format PDF) describing the main characteristics of their system. These contributions will be distributed to all the participants together with the final report of the organizing committee of the ALBAYZIN-08 TTS evaluation in the V Workshop on Speech Technology. In order to homegenize the aspect of the papers, please use the Latex or MS-WORD templates available in the JTH2008 web.

3. Each participant should agree to the free distribution of the synthetic speech signals among the rest of participants for research purposes.

4. To be present at the V Workshop on Speech Technology (JTH2008), to be held on November 12-14th 2008, to show the technical details of the TTS.