

## Post-doctoral position (12 months) - GIPSA-lab, Grenoble, France

### Incremental text-to-speech synthesis for people with communication disorders

#### Duration, location and staff

The position is open from January 2015 (until filled) for a duration of 12 months. The work will take place in GIPSA-lab, Grenoble, France, in the context of the project *SpeakRightNow*.

Researchers involved: Thomas Hueber, Gérard Bailly, Laurent Girin and Mael Pouget (PhD student).

#### Context

*SpeakRightNow* project aims at developing an incremental Text-To-Speech system (iTTS) in order to improve the user experience of people with communication disorders who use a TTS system in their daily life. Contrary to a conventional TTS, an iTTS system aims at delivering the synthetic voice while the user is typing (eventually with a delay of one word), and thus before the full sentence is available. By reducing the latency between text input and speech output, iTTS should enhance the interactivity of communication. Besides, iTTS could be chained with incremental speech recognition systems, in order to design highly responsive speech-to-speech conversion system (for application in automatic translation, silent speech interface, real-time enhancement of pathological voice, etc.).

The development of iTTS system is an emerging research field. Previous work mainly focused on the online estimation of the target prosody from partial (and uncertain) syntactic structure [2], and the reactive generation of the synthetic waveform (as in [3] for HMM-based speech synthesis). The goal of this post-doctoral position is to propose original solutions to these questions. Depending on the his/her background, the recruited researcher is expected to contribute on one or more of the following tasks:

- 1) Developing original approaches to address the problem of **incremental prosody estimation** using machine learning techniques for predicting missing syntactic information and driving prosodic models.
- 2) Implementing a prototype of an iTTS system on a **mobile platform**. The system will be adapted from the HMM-based TTS system currently developed at GIPSA-lab for French language.
- 3) Evaluating the prototype in a **clinical context** in collaboration with the medical partners of the *SpeakRightNow* project.

**Keywords:** assistive speech technology, incremental speech synthesis, prosody, machine learning, handicap.

**Prerequisite:** PhD degree in computer science, signal processing or machine learning. A background in HMM-based speech synthesis and/or development on iOS/Android platform is a plus.

**To apply:** Applicants should email a CV along with a brief letter outlining their research background, a list of two references and a copy of their two most important publications, to Thomas Hueber (thomas.hueber@gipsa-lab.fr).

#### References:

- [1] Baumann, T., Schlangen, D., "Evaluating prosodic processing for incremental speech synthesis," in Proceedings of Interspeech, Portland, USA, Sept. 2012.
  - [2] Astrinaki, M., d'Allessandro, N., Picart, B. Drugman, T., Dutoit, T., "Reactive and continuous control of HMM-based speech synthesis," in Proceedings of IEEE Workshop on Spoken Language Technology Miami, USA, Dec. 2012.
-