Call for Ph.D. thesis applications for a project entitled “Automatic recognition and generation of Cued Speech using deep learning”

RESEARCH FIELD: visual speech recognition, cued speech, machine learning, deep learning, multimodality, end-to-end model, sequence-to-sequence mapping

SUPERVISION: Dr. Denis BEAUTE MPS and Dr. Thomas HUEBER, CNRS researchers at GIPSA-lab, Grenoble, France

SCIENTIFIC DEPARTMENT GIPSA-lab, CNRS/Grenoble-Alpes University, Grenoble, France

DOCTORAL SCHOOL: Ecole doctorale Electronique-Electrotechnique-Automatique & Traitement du Signal (EEATS), Communauté Université Grenoble Alpes, Grenoble, France.

Envisioned start, duration and funding
The funding will cover the 3 years of a full-time PhD (2867.26 € monthly gross salary corresponding to 2304.41 € after employer and employee deductions and before individual tax in France), plus 600 € per month (submitted to employee deductions) as Mobility Allowance and if entitled a Family Allowance of 500 € per month (submitted to employee deductions). Ideally the starting date would be between September or October 2020. The ESR will also have access to funds to cover his/her research and training costs.

Job description
Cued speech (CS) is a gesture-based communication system used by deaf people. It uses a set of specific hand shapes and positions to complement the lip information and make all phonemes of a given spoken language clearly visible. The goal of this PhD project is two folds: 1) designing systems aiming at decoding automatically CS into text, 2) generating automatically realistic videos of a “virtual” CS interpreter from text. These modules are the necessary building block of a dialogue systems for deaf people.

This position is a part of the H2020 Marie-Curie Innovative Training Network project (Comm4CHILD) that received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 860755. Research will be conducted at GIPSA-lab in Grenoble, France in the CRISSP research team (Cognitive Robotics, Interactive Systems and Speech Processing, http://www.gipsa-lab.grenoble-inp.fr/en/crissp.php). A secondment will take place at the IVèS company (M12-M13 and M24-M25) and at ULB Université Libre de Bruxelles (M19-20). IVèS has developed a strong expertise in phone platform integrating end-users aspects and is very well identified at French national and international levels and especially in Grenoble, Toulouse (IVèS had recently acquired the ELIOZ company), and in Montreal. The expertise of the ULB in language development will enable the ESR to evaluate the different technical solutions in relation with language abilities of children with Hearing Implants.

Requested expertise
Candidates should have a strong background in one of these fields: computer vision, statistics, machine learning, natural language processing, speech processing - must have very good programming skills (in Python), and should have good verbal and written communication skills
in English. While not mandatory, basic knowledges in French will be appreciated since most of our datasets are based on French version of CS (called Langue française Parlée Complétée).

**Eligibility criteria:**
Candidates must have obtained a degree which formally entitles them to embark on a doctorate, either in the country in which the degree was obtained or in the country in which the researcher is recruited. The candidate must have resided or carried out their main activity (work or studies) in a different country from the host organization for at least 12 months in the last 3 years immediately before the recruitment date. Holidays are not counted. Candidates cannot have been awarded a doctoral degree and/or completed more than four years of full-time equivalent research experience.

**Applications:**
Interested candidates should submit (a) a cover letter describing their background, research experiences, interests, and goals, (b) a curriculum vitae, (c) at least one letter of recommendation from previous research supervisors, (d) a copy of diploma (with the university transcripts), to Denis.Beautemps@gipsa-lab.grenoble-inp.fr, Thomas.Hueber@gipsa-lab.grenoble-inp.fr (PhD supervisors) as well as to bjorn.lyxell@liu.se (recruitment officer of the grant). Please, mention that you are applying to the “DOCTORAL position Comm4CHILD” in the email subject.

**Scientific description**
The successful candidate will investigate advanced deep learning techniques to model the complex relationships between lips, hand gestures, and text in CS. The proposed work plan is the following: 1) extending existing datasets by recording new CS interpreters in both French and English using video and 3D motion capture technique, 2) investigating sequence-to-sequence mapping techniques (based for example on the transformer architecture and GANs) to decode CS into text, 3) investigated video generation techniques based on Conditional VAE or GAN to synthesize automatically realistic CS gestures from text. An extension to the automatic recognition and generation of signal language could be considered (in collaboration with the LIMSI lab, Orsay, France).

**Contact address**
Denis.Beautemps@gipsa-lab.grenoble-inp.fr