SECTOR: Higher Education Institution

LOCATION: France, Grenoble

RESEARCHER PROFILE:

- First stage researcher,

INSTITUTION: Univ. Grenoble Alpes, University of Innovation

One of the major research-intensive French universities, Univ. Grenoble Alpes\(^1\) enjoys an international reputation in many scientific fields, as confirmed by international rankings. It benefits from the implementation of major European instruments (ESRF, ILL, EMBL, IRAM, EMFL\(^2\)). The dynamic ecosystem, grounded on a close interaction between research, education and companies, has earned Grenoble to be ranked as the 5th most innovative city in the world. Surrounded by mountains, the campus benefits from a natural environment and a high quality of life and work environment. With 7000 foreign students and the annual visit of more than 8000 researchers from all over the world, Univ. Grenoble Alps is an internationally engaged university.

A personalized Welcome Center for international students, PhDs and researchers facilitates your arrival and installation.

In 2016, Univ. Grenoble Alpes was labeled «Initiative of Excellence ». This label aims at the emergence of around ten French world class research universities. By joining Univ. Grenoble Alpes, you have the opportunity to conduct world-class research, and to contribute to the social and economic challenges of the 21st century (“sustainable planet and society”, “health, well-being and technology”, “understanding and supporting innovation: culture, technology, organizations” “Digital technology”).

* ESRF (European Synchrotron Radiation Facility), ILL (Institut Laue-Langevin), IRAM (International Institute for Radio Astronomy), EMBL (European Molecular Biology Laboratory), EMFL (European Magnetic Field Laboratory)

Key figures:

- + 50,000 students including 7,000 international students
- 3,700 PhD students, 45% international
- 5,500 faculty members
- 180 different nationalities
- 1st city in France where it feels good to study and 5th city where it feels good to work
- ISSO: International Students & Scholars Office affiliated to EURAXESS

---

\(^1\) Univ. Grenoble Alpes
MANDATORY REFERENCES:

IDEX PROJECT TITLE: Multidisciplinary Institute for Artificial Intelligence – Speech chair (P. Perrier)
SUBJECT TITLE: Hybrid Bayesian and deep neural modeling for weakly supervised learning of sensory-motor speech representations
RESEARCH FIELD: Computer Science, Cognitive Science, Machine Learning, Artificial Intelligence, Speech Processing
SCIENTIFIC DEPARTMENT (LABORATORY'S NAME): GIPSA-lab
DOCTORAL SCHOOL’S: MSTII (maths and computer science) or EEATS (signal processing) or EDISCE (cognitive science), depending on the candidate’s profile and career plan
SUPERVISOR’S NAME: J. Diard (LPNC); T. Hueber, L. Girin, J.-L. Schwartz (GIPSA-Lab)

SUBJECT DESCRIPTION:

General objective
How can a child learn to speak from hearing sounds, without any motor instruction provided by his/her environment? The general objective of this PhD project is to develop a computational agent, able to learn speech representations from raw speech data in a weakly supervised configuration. This agent will involve an articulatory model of the human vocal tract, an articulatory-to-acoustic synthesis system, and a learning architecture combining deep learning algorithms and developmental principles inspired from cognitive sciences. This PhD will be part of the “Speech communication” chair in the Multidisciplinary Institute for Artificial Intelligence in Grenoble (MIAI).

Method
This work will capitalize on two bricks of research recently developed in Grenoble. First, a Bayesian computational model of speech communication called COSMO (Communicating about Objects using SensoriMotor Operations) (Moulin-Frier et al., 2012, 2015; Laurent et al., 2017; Barnaud et al., 2019) was jointly developed by GIPSA and LPNC. This model associates speech production and speech perception models in a single architecture. The random variables in COSMO represent the signals and the sensori-motor processes involved in the speech production/perception loop. COSMO learns their probability distributions from speech examples provided by the environment, and is then able to perceive and produce speech sounds associated to speech categories. So far, COSMO was mostly tested on synthetic data. One of the main challenges is now to confront COSMO to real-world data.

Second, we will also capitalize on a set of computational models for automatic processing and learning of sensory-motor distributions in speech developed at GIPSA. This comprises a set of transfer-learning algorithms (Hueber et al., 2015, Girin et al. 2017) aiming at adapting acoustic-articulatory knowledge on one speaker, towards another speaker, using a limited amount of data, possibly incomplete and noisy; together with a set of deep neural networks able to process raw articulatory data (Hueber et al., 2016; Tatulli & Hueber, 2017).

The first step will consist in designing, implementing and testing a “deep” version of COSMO, in which some of the probability distributions are implemented by generative neural models (e.g. VAE, GAN). This choice is motivated by the ability of such techniques to deal with raw, noisy and complex data, as well as their flexibility in terms of transfer learning. The second stage will consist in reformulating entirely the speech communication agent in an end-to-end neural architecture.

Outputs
The system will be tested in terms of both efficiency of the learning process – hence ability to generate realistic speech sequences after convergence – and coherence of the motor strategies discovered by the computational agent, in spite of the fact that no motor data will be provided for learning. The outputs are both (1) theoretical – for better understanding the cognitive processes at hand in speech development and speech communication; (2) technical – for integrating knowledge about speech production and cognitive processes in a machine learning architecture; and (3) technological – for proposing a new generation of autonomous speech technologies.
**ELIGIBILITY CRITERIA**

Applicants must have:
- A Master’s degree (or be about to earn one) or have a university degree equivalent to a European Master’s (5-year duration), in Computer Science, Cognitive Science, Signal Processing or Applied Mathematics.
- Solid skills in Machine Learning or probabilistic modeling + General knowledge in natural language processing and/or speech processing (an affinity for cognitive sciences and speech sciences is welcome).
- Very good programming skills (mostly in Python).
- Good oral and written communication in English.
- Ability to work autonomously and in collaboration with supervisors and other team members.

Applicants will have to send an application letter in English and attach:
- Their last diploma
- Their CV
- A short presentation of their scientific project (2 to 3 pages max)
- Letters of recommendation are welcome.
- Contact before preparing a complete application are welcome too.

Applications will be evaluated as they are received: the position is open until it is filled, with deadline on July 10th, 2019.

Address to send their application: Jean-Luc.Schwartz@gipsa-lab.fr, Thomas.Hueber@gipsa-lab.fr