Master 2 Research Internship Offer 2026

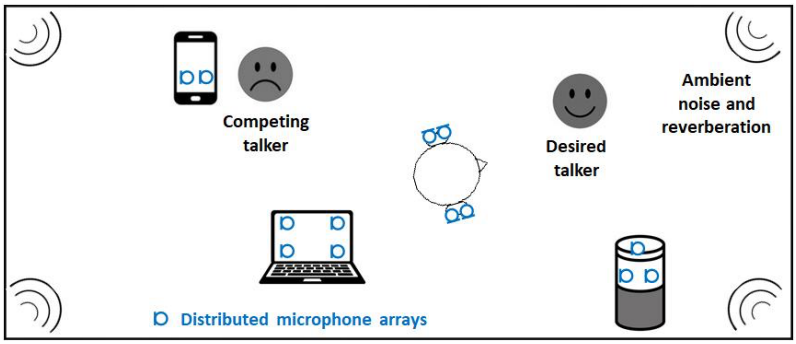
Speech Enhancement with a Distributed Microphone Array by Combining Acoustics and Machine Learning

Desired Date: Start in **March 2026**, for 5-6 months. **Contact:** antoine.deleforge@inria.fr

Required Skills

- Excellent level in **Python** programming. PyTorch knowledge is an added value.
- Training in **Deep Learning** and **Signal Processing**. Additional knowledge or interest for audio, acoustics, numerical methods or optimization are an added value.
- 2nd year master level (in computer science, signal processing, machine learning, acoustics or applied mathematics) with a **strong interest for academic research**.

Topic Description



Microphones are now ubiquitous in our environment: **hearing aids**, **smart speakers**, smartphones, **augmented reality** headsets, **videoconferencing** systems, etc. The comfortable use of these devices requires algorithms that **enhance speech signals of interest**, even in the presence of noise and reverberation. This task is made all the more difficult in dynamic conditions, where sources and microphones can move around. This internship is part of a French-German research project (ANR-DFG AWESOME 2026-2029) aimed at **exploiting all the microphones available** in a room, thus forming a *distributed* or *ad-hoc* array to greatly improve the quality of the speech signals captured.

The main difficulty in doing so is that the relative and absolute positions of the microphones in the room, as well as their acoustic characteristics and those of the room itself, are generally only **partially known**, which prevents the array from being exploited to its full potential. To overcome this obstacle, this internship will explore approaches combining **inverse acoustic methods** and **machine learning**, and in particular recent **diffusion-based generative models**. Two avenues could be explored:

- **Multichannel dereverberation preserving early reflections.** A dereverberation approach such as [1] or [2] will be extended to the scenario under consideration, then combined with an inverse acoustic method such as [3] to locate devices with respect to their nearest reflectors.
- **Calibration using hand claps.** A Schrödinger Bridge diffusion model [4] will be used to transform hand clap recordings into *room impulse responses*, the early parts of which can then be used by the inverse method in [5] to locate reflectors.

References (Please examine at least 3 of these references before applying – pdf under the links)

[1] E. Moliner, J-M. Lemercier, S. Welker, T. Gerkmann, and V. Välimäki. "BUDDy: Single-channel blind unsupervised dereverberation with diffusion models." In *18th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pp. 120-124. IEEE, 2024.

[2] L. Bahrman, M. Fontaine, and G. Richard. "A Hybrid Model for Weakly-Supervised Speech Dereverberation." In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1-5. IEEE, 2025.

[3] D. Di Carlo, C. Elvira, A. Deleforge, N. Bertin, and R. Gribonval. "BLASTER : An off-grid method for blind and regularized acoustic echoes retrieval." In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 156-160. IEEE, 2020.

[4] A. Jukić, R. Korostik, J. Balam, and B. Ginsburg. "Schrödinger Bridge for Generative Speech Enhancement." In *Proceedings of Interspeech*, pp. 1175-1179. ISCA. 2024.

[5] T. Sprunck, A. Deleforge, Y. Privat, and C. Foy. "Gridless 3D recovery of image sources from room impulse responses." *IEEE signal processing letters* 29 (2022): 2427-2431.

Practical Information

The internship will take place in **Strasbourg** (11 rue Jean-Mentelin) within a joint team between Inria and UMRAE and will be co-supervised by Antoine Deleforge (Inria researcher), Cédric Foy (UMRAE researcher), and Jean-Daniel Pascal (UMRAE-Inria PhD student). **The internship may lead to a fully-funded PhD thesis.**