Noureddine Khaous

(+33) 7 51 85 61 31
noureddine.khaous@etu.u-paris.fr
noureddinekhaous.com

PhD student in Natural Language and Speech Processing, specializing in Automatic Speech Recognition (ASR). My research focuses on improving speech models in multilingual and low-resource settings by understanding how they process and represent linguistic information, and using this knowledge to build better and more robust models.

Work Experience

- **Teaching Assistant** *Sep 2026 – Present*
  Université Paris Cité Paris, France
  - Teaching Introduction to Python, covering programming fundamentals and supervising hands-on practical sessions.

- **Data Scientist** *Sep 2023 – Sep 2025*
  Orange [🌐] Paris, France
  - Adapted large language models to Wolof through dataset construction, vocabulary expansion, and distributed fine-tuning with DeepSpeed, achieving a 12% METEOR improvement on Wolof→French and Wolof→English machine translation.
  - Built and deployed an Agentic RAG system with LangGraph for network incident management, integrating summarization, chatbot support, and next-action prediction, reducing ticket processing time by 50%.

- **Research Intern** *May 2025 – Jul 2025*
  Bergische Universität Wuppertal Remote
  - Collaborating on AI models integrating physics for fluid state prediction in nonlinear port-Hamiltonian systems.
  - Developing conditional diffusion models to generate spatially-resolved system state predictions in image form.
  - Supporting system identification and simulation of complex fluid dynamics processes.
  - Conducting research at the intersection of machine learning, physics-informed AI, and dynamical systems.

- **Research Intern** *Oct 2024 – Jan 2025*
  IDEMIA [🌐] Paris, France
  - Improved contactless fingerprint image quality using state-of-the-art diffusion models for biometric applications.
  - Designed an attention-based U-Net architecture with time-step encoding for image denoising.
  - Enhanced similarity between contactless and contact-based fingerprint images, improving recognition robustness.
  - Conducted experiments on generative image restoration methods for secure identity systems.

Education

- **PhD in Machine Learning for Low-Resource Languages** *Feb 2026 – Present*
  Université Paris Cité | *Laboratoire de Linguistique Formelle* Paris, France
  - Topic: Machine learning for rare languages: linking speech recognition and analysis of the internal representations of Speech Models in service of field linguistics | Supervisor: *Guillaume Wisniewski*

- **Master of Science in Computer Science and Robotics** *Sep 2023 – Sep 2025*
  Sorbonne Université Paris, France
  - Focus on Machine Learning, Speech Processing, Computer Vision and Robotics.
  - Dissertation Topic: Adapting Large Language Models to Low-Resource Languages: A Case Study of Wolof.
  - Graduated Rank: 3/45

- **Bachelor of Engineering in Electronics and Control Theory** *Sep 2021 – May 2023*
  Université Claude Bernard Lyon 1 Lyon, France
  - Focus on Electronics, Electrical Engineering, and Control Theory.
  - Graduated Rank: 4/50

- **Preparatory Classes for Engineering Schools | Mathematics and Physics Track** *Sep 2019 – Jun 2021*
  École Nationale Polytechnique Algiers, Algeria
  - Intensive coursework in Mathematics, Linear Algebra, Physics, Chemistry, and Engineering Sciences.
  - Graduated Rank: 19/283

PUBLICATIONS

- **Khaous, N.** & Wisniewski, G. (2025). *Accent-Induced Geometric Displacement in Speech Representations Predicts Automatic Speech Recognition Degradation*. *TALN 2026.*
- El Messaoudi, A., **Khaous, N.**, Cherifi, K. (2026). *Target Parameterization in Diffusion Models for Nonlinear Spatiotemporal System Identification*. *arXiv preprint 2604.17566.*

SKILLS

- **Programming Languages:** Python, C/C++, SQL, Bash
- **Machine Learning & AI:** PyTorch, Hugging Face, Deep Learning, NLP, LLMs, Transformers, Computer Vision
- **Data Science & Scientific Computing:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn
- **Tools & Technologies:** Docker, Git, Linux, OpenCV, LaTeX
- **Research Areas:** Speech Processing, Representation Learning, Low-Resource Languages, Multilingual NLP
- **Languages:** French (Native), Arabic (Native), Kabyle (Native), English (Full Professional Proficiency)