Martin Kocour

martinkocour1102@gmail.com | MartinKocour | martinkocour | Prague, Czechia

EXPERIENCE

Brno University of Technology, Speech@FIT Sep. 2019 – Present
PhD. Student and Junior Researcher Brno, Czechia

- Research on multi-talker speech recognition and speaker diarization [2], [4]
- Research on noise robust automatic speech recognition for air-traffic management, including acoustic modelling and contextual language model adaptation; mainly for Horizon 2020 ATCO^2 project [8], [9]
- Supervising projects, teaching numerical and computer exercises in several courses: Signals and Systems, Bayesian Models for Machine Learning and Machine Learning and Recognition
- PhD. supervisor: prof. Dr. Ján "Honza" Černocký

Apple Mar. 2024 – Aug. 2024
Research Intern Boston, MA, USA

- Developed a multi-modal speech recognition system that integrated next-token predictions from LLM to enable dynamic contextual adaptation in an encoder-decoder ASR architecture, improving recognition performance
- Working on techniques to minimize latency of inference with large language models [1]
- Research supervisors: Adnan Haider, PhD, Takaaki Hori, PhD and Zhen Huang, PhD

JSALT - Frederick Jelinek Summer Workshop Jun. 2023 – Aug. 2023
Visiting researcher Le Mans, France

- Design and development of modern WFST library supporting automatic differentiation
- Research supervisor: Lucas Ondel Yang, PhD

Mila – Quebec Artificial Intelligence Institute Oct. 2022 – Jan. 2023
Visiting researcher Montreal, Canada

- Research on joint multi-talker speech recognition and speaker diarization
- Contribution to SpeechBrain – An Open-Source Conversational AI Toolkit
- Research supervisor: Mirco Ravanelli, PhD

Parrot Sep. 2019 – Oct. 2023
Machine Learning Engineer Prague, Czechia

- Built a Kaldi-based ASR system for the U.S. court reporting service
- Developed a WFST-based force-alignment system to generate word-level timestamps for transcriptions, enabling efficient search and indexing within reports
- Experience with hybrid ASR systems and pre-trained end-to-end ASR systems fine-tuned on internal datasets

Telefónica R&D June 2018 – Sept. 2018
Research Intern Barcelona, Spain

- Design and development of light-weight ASR and object detection systems and their deployment into 5G antenna (BTS), i.e., a device with constrained computational resources [12]
- Lightly-supervised training of Spanish ASR system based on subtitled speech and experiments with incremental lightly-supervised learning [13]
- Research supervisor: Jordi Luque, PhD

Dotykačka s.r.o May 2015 – May 2018
Android Developer Brno, Czechia

- Development of Dotykačka app, a mobile "Point of Sale" system for retail and restaurants
- App deployment, auto-scaling of custom services, and administration of virtual machines

EDUCATION

Brno University of Technology Sep. 2017 – June 2019
Master Degree in Information Technology Brno, Czechia

- **Relevant coursework**: Signal Processing, Data Structures, Algorithms, Computer Systems, Machine Learning
- **Master thesis:** Automatic Speech Recognition System Continually Improving Based on Subtitled Speech Data, co-supervised by Jordi Luque (Telefónica R&D) and Honza Černocký (BUT) [13]

SKILLS

Programming : Python, Julia, C++, Bash
Frameworks: Hugging Face, k2, Kaldi
Tools: PyTorch, Lhotse, Jax, NumPy, Git
Languages : English, Slovak, German (beginner)

AWARDS

- **Winner** of the Albayzín's 2022 Speech-to-Text Challenge [3], [6] 2022
- Placed **3rd** in Albayzín's 2020 Speech-to-Text Challenge, evaluating ASR systems for TV shows [7] 2021
- Placed **5th** in the CHiME 6 (track2), a competition on conversational speech recognition [11] 2020
- Received **Dean's award** for outstanding Master's thesis 2019

PUBLICATIONS

[1] T. Hori, **M. Kocour**, A. Haider, E. McDermott, and X. Zhuang, "Delayed Fusion: Integrating Large Language Models into First-Pass Decoding in End-to-end Speech Recognition," in *IEEE ICASSP*, 2025.

[2] A. Polok*, D. Klement*, **M. Kocour***, J. Han, F. Landini, B. Yusuf, M. Wiesner, S. Khudanpur, J. Cernocký, and L. Burget, "DiCoW: Diarization-Conditioned Whisper for Target Speaker Automatic Speech Recognition," *Submitted to Computer Speech & Language*, 2025.

[3] K. Benes, **M. Kocour**, and L. Burget, "Hystoc: Obtaining Word Confidences for Fusion of End-To-End ASR Systems," in *IEEE ICASSP*, 2024.

[4] **M. Kocour**, K. Zmolikova, L. Ondel, J. Svec, M. Delcroix, T. Ochiai, L. Burget, and J. Cernocký, "Revisiting joint decoding based multi-talker speech recognition with DNN acoustic model," in *InterSpeech*, 2022.

[5] L. Ondel, L.-M. Lam-Yee-Mui, **M. Kocour**, C. F. Corro, and L. Burget, "GPU-Accelerated Forward-Backward Algorithm with Application to Lattice-Free MMI," in *IEEE ICASSP*, 2022.

[6] J. Umesh*, **M. Kocour***, M. Karafiat*, J. Svec*, F. López*, K. Benes*, *et al.*, "BCN2BRNO: ASR System Fusion for Albayzin 2022 Speech to Text Challenge," in *IberSPEECH*, 2022.

[7] **M. Kocour**, G. Cámbara, J. Luque, D. Bonet, M. Farrús, M. Karafiát, K. Veselý, and J. Cernocký, "BCN2BRNO: ASR System Fusion for Albayzin 2020 Speech to Text Challenge," in *IberSPEECH*, 2021.

[8] **M. Kocour**, K. Veselý, A. Blatt, J. Z. Gomez, I. Szöke, J. Cernocký, D. Klakow, and P. Motlícek, "Boosting of Contextual Information in ASR for Air-Traffic Call-Sign Recognition," in *InterSpeech*, 2021.

[9] **M. Kocour**, K. Veselý, I. Szöke, S. Kesiraju, J. Zuluaga-Gomez, A. Blatt, *et al.*, "Automatic processing pipeline for collecting and annotating air-traffic voice communication data," in *OpenSky Symposium*, 2021.

[10] I. Szöke, S. Kesiraju, O. Novotný, **M. Kocour**, K. Veselý, and J. Cernocký, "Detecting English Speech in the Air Traffic Control Voice Communication," in *InterSpeech*, 2021.

[11] K. Zmolíková, **M. Kocour***, N. F. Landini*, K. Benes*, M. Karafiát*, K. H. Vydana, *et al.*, "BUT System for CHiME-6 Challenge," in *CHiME Virtual Workshop*, 2020.

[12] A. Cartas, **M. Kocour**, A. Raman, I. Leontiadis, J. Luque, N. Sastry, L. Nunez-Martinez, D. Perino, and S. C. Perales, "A reality check on inference at mobile networks edge," in *EDGESYS*, 2019.

[13] **M. Kocour**, "Automatic Speech Recognition System continually improving based on subtitled speech data," M.S. thesis, Brno University of Technology, Faculty of Information Technology, Brno, 2019.

*These authors contributed equally.