Martin Kocour

(+420)721430354 | martinkocour1102@gmail.com | martinkocour | Prague, Czechia

EXPERIENCE

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

- Research on decoding techniques for single channel multi-talker speech recognition (cocktail party problem) [10],
- 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 [7, 8],
- Supervising projects, teaching numerical and computer exercises in several courses: Signals and Systems, Bayesian Models for Machine Learning and Machine Learning and Recognition,
- Participation in Chime6 Challenge and Albayzin 2020 and 2022 Speech-to-Text Challenge (winning system) [1, 6, 9],
- PhD. supervisor: Honza Černocký.

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

- Leveraging language models (OpenLLaMA 7B, Mistral 7B) to correct ASR transcription errors.
- Working on techniques to minimize latency of inference with large language models [3].
- Research supervisors: Adnan Haider, PhD & Takaaki Hori, PhD.

JSALT - Frederick Jelinek Summer Workshop Jun. 2022 – 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

- Development of LVCSR system for the U.S. court reporting service.
- Working with both: a) hybrid ASR systems in Kaldi (e.g. CNN-TDNNf acoustic model) and b) E2E ASR systems – pre-trained models from Hugging Face (e.g. wav2vec 2.0) fine-tuned on internal dataset with MMI objective.

Telefónica R&D June 2018 – Sept. 2018
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 [2],
- Lightly-supervised training of Spanish ASR system based on subtitled speech and experiments with incremental lightly-supervised learning [5],
- 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, automation, continuous integration, auto-scaling 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, Databases, 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) [5],
- Dean's award for excellent Master Thesis.

SKILLS

Computer : Programming in C++, Python, Julia. Working with Kaldi, PyTorch, GitHub, Weights & Biases.

Languages : English (fluent), German (B1), Slovak (mother tongue).

INTERESTS

Hiking, White water rafting, Climbing, Basketball, Ski mountaineering, Music, Friends.

PUBLICATIONS

[1] Karel Beneš, **Martin Kocour**, and Lukáš Burget. Hystoc: Obtaining word confidences for fusion of end-to-end asr systems. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 11276–11280, 2024. doi: 10.1109/ICASSP48485.2024.10446739.

[2] Alejandro Cartas, **Martin Kocour**, Aravindh Raman, Ilias Leontiadis, Jordi Luque, Nishanth Sastry, Leon Nunez-Martinez, Diego Perino, and Segura Carlos Perales. A reality check on inference at mobile networks edge. In *Proceedings of the 2nd ACM International Workshop on Edge Systems, Analytics and Networking (EDGESYS '19)*, pages 54–59. Association for Computing Machinery, 2019. ISBN 978-1-4503-6275-7. doi: 10.1145/3301418.3313946. URL https://www.fit.vut.cz/research/publication/11956.

[3] Takaaki Hori, **Martin Kocour**, Adnan Haider, and Erik McDermott. Delayed fusion: Integrating large language models into first-pass decoding in end-to-end speech recognition. In *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025.

[4] Lucas Ondel, Léa-Marie Lam-Yee-Mui, **Martin Kocour**, Caio Filippo Corro, and Lukás Burget. Gpu-accelerated forward-backward algorithm with application to lattice-free mmi. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8417–8421, 2022. doi: 10.1109/ICASSP43922.2022.9746824.

[5] **Martin Kocour**. Automatic Speech Recognition System Continually Improving Based on Subtitled Speech Data. Master's thesis, Brno University of Technology, Faculty of Information Technology, Brno, 2019.

[6] **Martin Kocour**, Guillermo Cámbara, Jordi Luque, David Bonet, Mireia Farrús, Martin Karafiát, Karel Veselý, and Jan Černocký. BCN2BRNO: ASR system fusion for Albayzin 2020 speech to text challenge. In *Proceedings of IberSPEECH 2021*, pages 113–117. International Speech Communication Association, 2021. doi: 10.21437/IberSPEECH.2021-24. URL https://www.fit.vut.cz/research/publication/12577.

[7] **Martin Kocour**, Karel Veselý, Alexander Blatt, Juan Zuluaga Gomez, Igor Szöke, Jan ernocký, Dietrich Klakow, and Petr Motlicek. Boosting of contextual information in ASR for air-traffic call-sign recognition. In *Proc. Interspeech 2021*, pages 3301–3305, 2021. doi: 10.21437/Interspeech.2021-1619.

[8] **Martin Kocour**, Karel Veselý, Igor Szőke, Santosh Kesiraju, Juan Zuluaga-Gomez, Alexander Blatt, Petr Motlíček, et al. Automatic processing pipeline for collecting and annotating air-traffic voice communication data. In *Proceedings of the 9th OpenSky Symposium 2021*. MDPI, 2021.

[9] **Martin Kocour**, Jahnavi Umesh, Martin Karafiat, Ján vec, Fernando López, Jordi Luque, Karel Bene, Mireia Diez, Igor Szoke, Karel Veselý, Luká Burget, and Jan ernocký. Bcn2brno: Asr system fusion for albayzin 2022 speech to text challenge. In *IberSPEECH 2022*, pages 276–280, 2022. doi: 10.21437/IberSPEECH.2022-56.

[10] **Martin Kocour**, Katerina Zmolikova, Lucas Ondel, Jan Svec, Marc Delcroix, Tsubasa Ochiai, Lukas Burget, and Jan Cernocky. Revisiting joint decoding based multi-talker speech recognition with DNN acoustic model. In *Proc. Interspeech 2022*, pages 4955–4959, 2022. doi: 10.21437/Interspeech.2022-10406. URL https://arxiv.org/abs/2111.00009.

[11] Kateřina Žmolíková, **Martin Kocour**, Nicolás Federico Landini, Karel Beneš, Martin Karafiát, K. Hari Vydana, Alicia Díez Lozano, Oldřich Plchot, K. Murali Baskar, Ján Švec, Ladislav Mošner, Vladimír Malenovský, Lukáš Burget, Bolaji Yusuf, Ondřej Novotný, František Grézl, Igor Szőke, and Jan Černocký. BUT system for CHiME-6 challenge. In *Proceedings of CHiME 2020 Virtual Workshop*, pages 1–3. University of Sheffield, 2020. doi: 10.21437/CHiME.2020-13. URL https://www.fit.vut.cz/research/publication/12283.