Roman Koshkin

1919-1, Tancha, Onna-son, Kunigami-gun, Okinawa 904-0495, Japan

https://roman-koshkin.unit.oist.jp roman.koshkin@gmail.com @romankoshkininterpreter https://github.com/RomanKoshkin/

❖ Final-year PhD student in machine learning and computational neuroscience. ❖ Experienced in simultaneous machine translation and language modeling ❖ Interested in natural language processing and understanding.

EDUCATION

09/2019 - present **Okinawa Institute of Science and Technology,** *Neural Coding and Brain Computing Unit, Japan*
PhD (Machine Learning and Computational Neuroscience) *Expected graduation*: **06/2025**

06/2017 **National Research University HSE,** *Moscow, Russia*
Master of Science, Psychology (**with distinction**, GPA: 8.9/10)

06/2002 **VUMO University,** *Moscow, Russia*
Specialist, Linguistics (**with honors**, GPA: 4.9/5)

WORK EXPERIENCE

07/2023 - 10/2023 **Special Research Intern**, ***NLP Group**, AHC Lab, Nara Institute of Science and Technology, Japan*
- Developed speech-to-text and speech-to-speech SiMT models leveraging open-source causal LLMs.
- Set up LLMOps/MLOps, parallel experiments to identify best design and HP choices.

07/2022 - 11/2022 **Research Intern**, ***Araya**, Reinforcement Learning Research Team, Tokyo, Japan*
- Conduct research towards using EEG for robot control with a brain-machine interface
- Compiled a sensor-aligned motor imagery EEG dataset on which I
- Trained a self-supervised EEG feature extractor with a contrastive loss and
- Achieved competitive performance in downstream tasks (incl. MI imagery classification).
- Reimplemented and open-sourced an M/EEG speech-decoding model.

09/2017 - 07/2019 **Junior Research Fellow**, ***Center for Bioelectric Interfaces**, Institute of Cognitive Neuroscience, National Research University HSE, Moscow, Russia*
- Coordinated a research team of 3 people for 2 years
- Conceptualized and conducted neuromarketing and consumer behavior research experiments
- Designed and implemented EEG data collection and pre-processing pipelines
- Wrote and maintained data acquisition software (Python front- & backend)
- Designed and trained DL models for estimating respondents' opinion of advertised products
- Taught EEG data pre-processing techniques, Python and MATLAB to junior lab members
- Provided oral status updates and written progress reports to the funding company (Neurotrend)
- Co-authored one patent (RF Patent 2747571)

PROJECTS

06/2024 - 04/2024 **Multilingual zero-shot simultaneous machine translation**
Neural Coding and Brain Computing Unit, OIST
LLM-based speech-to-text simultaneous machine translation with no costly pre-training or fine-tuning.

07/2023 - 10/2023 **TransLLaMa**
NLP Group, NAIST
LLM-based speech-to-text simultaneous machine translation.

01/2023 - 03/2023 **convSeq**
Neural Coding and Brain Computing Unit, OIST
Fast and scalable convolution-based method for unsupervised detection of patterns in neural recordings.

09/2020 - present **SoNNet**
Neural Coding and Brain Computing Unit, OIST
High-performance C++ library with a configurable user-friendly Python API for building recurrent spiking neural networks (SNNs).

09/2022 - 12/2022 **graphSeq**
Neural Coding and Brain Computing Unit, OIST
Graph neural network-based method for embedding and clustering of neural spiking patterns.

09/2022 - 12/2022 **M/EEG-based zero-shot speech decoding**
Araya Lab, Tokyo
Re-implementation of an algorithm that decodes speech from human brain recordings (M/EEG) 0-shot.

04/2020 - 09/2020 **Tutoring Object Manipulation Skills in a Human-Robot Interaction Paradigm**
OIST Cognitive Neurorobotics Unit, OIST
Trained a robot to perform reach-and-grasp tasks by combining learned motor primitives.

1/2020 - 4/2020 **Backpropagation-free learning for classification tasks**
OIST Neural Coding and Brain Computing Unit
Built a spike-timing dependent plasticity-based spiking neural network for image classification.

09/2019 - 12/2019 **Extended Ca^{2+} Buffer and Dynamics Model of the Rat Hippocampal Presynapse**
OIST Computational Neuroscience Unit
Implemented a reaction-diffusion model of Ca2+ dynamics in the rat hippocampal presynapse.

10/2017 - 07/2019 **Neurobarometer,** *Center for Bioelectric Interfaces, Higher School of Economics*
Software & algorithm for EEG-based neuromarketing and consumer behavior research.

09/2016 **Finding Weak Effects with Known Temporal Structure in Evoked Response Data,** *NRU HSE*
04/2017 Contributed to designing a novel projection-based method for identifying weak effects in noisy ERP data
09/2015 **Attention and Working Memory in Simultaneous Interpreting,** *Higher School of Economics*
09/2016 Tested the Efforts Model of simultaneous interpreting using the ERP technique

SKILLS

Frameworks/tools: *Pytorch (highest proficiency), HuggingFace, Lightning, scikit-learn, JAX, Apache Spark*
Infrastructure/HPC: *AWS, slurm*
Programming languages: *Python (highest proficiency), C++, Matlab, R, HTML, JavaScript*
Frontend development: *React, Next.js*
Virtualization tools: *Docker, Singularity*
Databases: *Neo4j, Redis, MongoDB*
LLMOps: *LangSmith, LangChain, LangGraph, wandb*

PATENTS

RF Patent 2747571. Electroencephalographic method and system of objective estimation of listeners' reaction to audio content based on a range of voluntary affective categories. https://bit.ly/EEGpatent2

AWARDS, GRANTS, FELLOWSHIPS

2023 **KAKENHI Grant-in-Aid (¥ 1.8M)** (https://cir.nii.ac.jp/crid/1040577431243576704)
2023 **Japan Society for the Promotion of Science Fellowship** (http://bit.ly/3PjzL7y)
2021 **Google PhD Fellowship ($ 10K)** (https://research.google/outreach/phd-fellowship/recipients/?category=2021)

PEER-REVIEWED PUBLICATIONS & PREPRINTS

Koshkin, R., Sudoh, K., Nakamura, S. (2024). LLMs Are Zero-Shot Context-Aware Simultaneous Translators. *arXiv.* https://arxiv.org/abs/2406.13476

Koshkin, R., Sudoh, K., Nakamura, S. (2024). TransLLaMa: LLM-based Simultaneous Translation System. *arXiv.* https://arxiv.org/abs/2402.04636

Koshkin, R., Fukai, T. (2024). convSeq: Fast and Scalable Method for Detecting Patterns in Spike Data. *ICML 2024* https://arxiv.org/abs/2402.01130

Koshkin, R., Fukai, T. (2023). Unsupervised Detection of Cell Assemblies with Graph Neural Networks. *In ICLR 2023 Tiny Papers Track.* https://openreview.net/pdf?id=Tbzv_BbjjO8

Koshkin, R., Shtyrov, Y. , Myachykov, A, & Ossadtchi, A. (2018). Testing the Efforts Model of Simultaneous Interpreting. *PLoS ONE 13*(10): e0206129. https://doi.org/10.1371/ journal.pone.0206129

Koshkin, R., & Ossadtchi, A. (2017). Commentary: Functional Connectivity in the Left Dorsal Stream Facilitates Simultaneous Language Translation: An EEG Study. *Frontiers in Human Neuroscience, 11*(2), 273. http://doi.org/10.3389/fnhum.2017.00064

Koshkin, R., Ossadtchi, A. & Shtyrov, Y. (2017). Attention, Working Memory And Listening In Simultaneous Interpreting. *Russian Journal of Cognitive Science, 4(4).* http://cogjournal.org/eng/4/4/index.html

Koshkin R. (2016). Comparative Analysis of Quantitative Dynamics of English-Russian and Russian-English Simultaneous Interpreting. *Bulletin of Moscow University, Series 22: Theory of Translation. Vol. 2,* 28-43 https://elibrary.ru/item.asp?id=27125259

POSTER PRESENTATIONS AND TALKS

Koshkin, R, Fukai, T. (2022). Astrocytes facilitate self-organization and remodeling of cell assemblies under STP-coupled STDP. SfN Conference, Nov 14-16, San Diego. Abstract: https://bit.ly/SfN_nov_2022

Koshkin, R., Fukai, T (2021). Leveraging Self-organized Structure for Memory Encoding in Binary Networks. RIKEN-OIST Symposium, Oct. 6-7, 2021, Japan Poster: https://bit.ly/3lgsqGO

Koshkin, R., Shtyrov, Y. & Ossadtchi, A. (2017). Testing One Aspect of the Efforts Model of Simultaneous Interpreting: An ERP Study. In *Proceedings of the Workshop "Neurobiology Of Speech And Language"*, Oct. 27-29, 2017, SPb, Russia Abstract: http://bit.ly/2y52Hu3 Poster: http://bit.ly/2ljEytV

Koshkin, R., Ossadtchi, A. & Shtyrov, Y.(2016). N1 ERP As an Index of Depth of Processing In Simultaneous Interpreting. In *Proceedings of Communication, Computation, and Cognitive Processes,* Sept. 28-29, 2016, Moscow, Russia Abstract: http://bit.ly/2lhyWjP

Koshkin, R., Ossadtchi, A. & Shtyrov, Y.(2017). Working Memory Load In Simultaneous Language Interpretation: An ERP Study. *IEEE International Symposium «Video and Audio Signal Processing in the Context of Neurotechnologies»*, Jun. 26-30, 2017, SPb, Russia Abstract: http://bit.ly/2ANhSVD

Kuznetsova A., **Koshkin R.**, Ossadtchi A. (2017). Localizing Hidden Regularities With Known Temporal Structure in the EEG Evoked Response Data. *IEEE International Symposium «Video and Audio Signal Processing in the Context of Neurotechnologies»*, Jun. 26-30, 2017, SPb, Russia Abstract: http://bit.ly/2ANhSVD

CONFERENCE PROCEEDINGS AND BOOK CHAPTERS

Koshkin, R., Ossadtchi, A. (2017). Working Memory Load in Simultaneous Language Interpretation: An ERP Study. In *Proc. of the 4th Conference "Cognitive Science in Moscow: New Research"*. July 15, 2017, Moscow, Russia. p. 434 http://virtualcoglab.ru/MoscowCogSci2017Proceedings.pdf

Garcia, A., **Koshkin, R.**, Paiva, T. (2023). *EEG In Cognitive Translation and Interpreting Studies.* JBPH. (In review)

SERVICE AND MENTORSHIP

03/2023 **Reviewer**, TPT @ ICLR 2023
11/2021 **Science Mentor**, Introduction to Deep Learning with Python, *Okinawa, Japan*