

Hosein Mohebbi

✉ h.mohebbi@tilburguniversity.edu | in hosein-mohebbi | t hmohebbi75 | hmohebbi.github.io
Tilburg, The Netherlands

Education & Research

Tilburg University | Ph.D. candidate

Nov 2021 – 2025 (expected)

I'm part of the consortium project [InDeep: Interpreting Deep Learning Models for Text and Sound](#)

Supervisors: Afra Alishahi (Tilburg University), Willem Zuidema (University of Amsterdam), and Grzegorz Chrupała (Tilburg University)

University of Edinburgh | Visiting researcher

Jan – Apr 2024

Visited ILCC, School of Informatics, University of Edinburgh, worked with Ivan Titov

Iran University of Science and Technology | Master's in Artificial Intelligence and Robotics

Sep 2019 – Sep 2021

Master's Thesis: "Interpretability and Transferability of Linguistic Knowledge in Pre-trained Language Models"

Supervisor: Mohammad Taher Pilehvar

Ferdowsi University of Mashhad | Bachelor's in Computer Engineering

Sep 2014 – Sep 2019

Bachelor's Thesis: "Improving 2D Soccer Agent's Decision Making using Deep Reinforcement Learning"

Supervisor: Ahad Harati

Publications

For the complete list please check out: [Google Scholar](#) | [Semantic Scholar](#) | [ACL Anthology](#) | [ORCID](#)

Disentangling Textual and Acoustic Features of Neural Speech Representations

Under review

Hosein Mohebbi • Grzegorz Chrupała • Willem Zuidema • Afra Alishahi • Ivan Titov | [\[paper\]](#) [\[code\]](#)

How Language Models Prioritize Contextual Grammatical Cues?

BlackboxNLP 2024

Hamidreza Amirzadeh • Afra Alishahi • Hosein Mohebbi | [\[paper\]](#)

DecoderLens: Layerwise Interpretation of Encoder-Decoder Transformers

Findings of NAACL 2024

Anna Langedijk • Hosein Mohebbi • Gabriele Sarti • Willem Zuidema • Jaap Jumelet | [\[paper\]](#)

Homophone Disambiguation Reveals Patterns of Context Mixing in Speech Transformers

EMNLP 2023

Outstanding Paper Award

Hosein Mohebbi • Grzegorz Chrupała • Willem Zuidema • Afra Alishahi | [\[paper\]](#) [\[code\]](#)

Quantifying Context Mixing in Transformers

EACL 2023

Hosein Mohebbi • Willem Zuidema • Grzegorz Chrupała • Afra Alishahi | [\[paper\]](#) [\[code\]](#) [\[blog\]](#) [\[demo\]](#)

AdapLeR: Speeding up Inference by Adaptive Length Reduction

ACL 2022

Ali Modarressi • Hosein Mohebbi • Mohammad Taher Pilehvar | [\[paper\]](#) [\[blog\]](#)

Not All Models Localize Linguistic Knowledge in the Same Place: A Layer-wise Probing on BERToids' Representations

BlackboxNLP 2021

Mohsen Fayyaz • Ehsan Aghazadeh • Ali Modarressi • Hosein Mohebbi • Mohammad Taher Pilehvar | [\[paper\]](#) [\[blog\]](#)

Exploring the Role of BERT Token Representations to Explain Sentence Probing Results

EMNLP 2021

Hosein Mohebbi • Ali Modarressi • Mohammad Taher Pilehvar | [\[paper\]](#) [\[blog\]](#)

Services

BlackBoxNLP 2023-2025 | Co-organizer

A popular Workshop on Analysing and Interpreting Neural Networks for NLP

Reviewing

I reviewed for the following conferences: EMNLP (2022, 2023), ACL 2023, EACL 2023, ACL Rolling Review (since 2022)

Industry Experience

Sadad Informatic Corporation - Elenoon | Software Developer
Summer Internship, 2017

Awards and Honors

- Got an Outstanding Paper Award at EMNLP 2023 Dec 2023
- Ranked 4th in Sharif Data Days National Competition, Sharif University of Technology Mar 2020
- Ranked 5th in RoboCup Asia Pacific International Robotics Competition, 2D Soccer Simulation, Kish Island Dec 2018
- Ranked 1st in Portuguese International Robotics Open (DE ROBOTICA), 2D Soccer Simulation, University of Coimbra Apr 2017
- Ranked 2nd in IranOpen International Robotics Competition, 2D Soccer Simulation League Apr 2017

Teaching Experience

Lecturer

- **Machine Learning** , co-lectured with Dr. Grzegorz Chrupala Tilburg University, Winter 2023, 2024
- **Methods for Responsible AI** , co-lectured with Dr. Seyed Mostafa Kia Tilburg University, Winter 2023, 2024

Teaching Assistant

- **Natural Language Processing** , lectured by Dr. Mohammad Taher Pilehvar [course webpage] Khatam University, Winter 2021
- **Artificial Intelligence** , lectured by Dr. Ahad Harati Ferdowsi University, 2017-2019
- **Computer Architecture** , lectured by Dr. Hamid Noori Ferdowsi University, 2016-2019
- **Basics of Computer Programming and Algorithms** , lectured by Dr. Saeid Abrishami Ferdowsi University, 2016-2019

Students Supervised

- Hamidreza Amirzadeh. MS. Topic: Interpretability .
- Mohsen Fayyaz, Ehsan Aghazadeh. BS. Topic: Interpretability.

Tutorials

- “Transformer-specific Interpretability”
The 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2024) [materials]
Hosein Mohebbi • Jaap Jumelet • Michael Hanna • Afra Alishahi • Willem Zuidema

Selected Talks

- Invited talk at CardiffNLP, Cardiff University Feb 2024
- Invited talk at the Centre for Speech Technology Research (CSTR), University of Edinburgh Feb 2024
- Invited talk on at GroNLP, University of Groningen Jun 2023
- Gave a guest lecture on Transformers to an undergraduate CL course at Tilburg University May 2023
- Invited talk at Cambridge/Cardiff Workshop in Natural Language Processing Jun 2021
- Gave a guest lecture on Interpretability to a graduate NLP course at Khatam University [slides] May 2021

References

Afra Alishahi a.alishahi@tilburguniversity.edu | [homepage]

Professor of Computational Linguistics at the Department of Cognitive Science and Artificial Intelligence at Tilburg University

Willem Zuidema w.h.zuidema@uva.nl | [homepage]

Associate professor at the Institute for Logic, Language and Computation, University of Amsterdam

Grzegorz Chrupala grzegorz@chrupala.me | [homepage]

Associate Professor at the Department of Cognitive Science and Artificial Intelligence at Tilburg University

Mohammad Taher Pilehvar mp792@cam.ac.uk | [homepage]

Senior Lecturer in the School of Computer Science at Cardiff University and an Affiliated Lecturer at the University of Cambridge

Ivan Titov ititov@inf.ed.ac.uk | [homepage]

Professor and Personal Chair in Natural Language Processing, ILCC, School of Informatics, University of Edinburgh