

Neslihan Cesur

🌐 neslihancesurr — [in neslihan-cesur](#)
🌐 neslihancesurr.github.io

✉ neslihan.cesur@ozu.edu.tr

☎ +90 538 086 68 69

SUMMARY

I am an M.Sc. student in Artificial Intelligence with a background in Linguistics, working to bridge the gap between theoretical structure and computational implementation. My early research focused on building foundational linguistic resources for Turkish, including Universal Dependencies and PropBank, as well as documenting endangered languages like Pazar Laz. Currently, I am focused on structure induction, where I develop rule-based algorithms to automatically generate Abstract Meaning Representation (AMR) graphs. For my future research, I hope to work on cognitively inspired deep learning architectures by integrating these structured representations into neural models, ensuring data efficiency and interpretability for low-resource languages.

EDUCATION

- 2023 – present M.Sc. in Artificial Intelligence, **Özyeğin University** (GPA: 3.47/4.0)
- 2019 – 2023 M.A. in Linguistics, **Boğaziçi University** (GPA: 4.0/4.0)
◊ Thesis: “Negation in Pazar Laz” (2023)
- 2013 – 2019 B.A. in Comparative Linguistics and Applied Foreign Languages, **Galatasaray University** (GPA: 3.8/4.0, graduated with honors, ranked 1st in department)
◊ Final dissertation: Collected spontaneous speech data and analyzed corpus for word order flexibility in Turkish
◊ 3rd year studies: Université Paris V Descartes, Department of *Sciences du Langage*

TEACHING AND RESEARCH EXPERIENCE

Research and Teaching Assistant, Özyeğin University August 2024 – present

Teaching assistant roles: grading homework and exams, proctoring, and running weekly lab sessions.

- ◊ CS410 Automata Theory and Formal Languages (Fall 2025-2026)
- ◊ CS452 Data Science with Python (Fall 2025-2026)
- ◊ CS201 Data Structures and Algorithms (Fall 2024-2025, Spring 2024-2025)
- ◊ MATH courses (Summer 2025)

Research and Teaching Assistant, Galatasaray University Jan 2021 - August 2024

Provided administrative and teaching support, including planning course schedules, proctoring exams, preparing school presentations, and delivering lectures. Developed and taught a preparatory course on linguistic typology and diversity for three consecutive years.

- ◊ Comparative Language Studies I (Fall 2021-2022, Fall 2022-2023, Fall 2023-2024)
- ◊ Comparative Language Studies II (Spring 2021-2022, Spring 2022-2023, Spring 2023-2024)

Linguist at StarLang Software Jan 2020 - Jan 2021

Responsible for defining the Universal Dependencies annotation scheme for Turkish and annotating multiple UD treebanks, including Atis (English), Turkish Penn TreeBank, FrameNet, Kenet, and Tourism. Collaborated on annotations to improve search engine performance for e-commerce platforms.

ONGOING PROJECTS

Project: Learning Abstract Meaning Representation for Turkish and Building a Question Answering System

Project Manager: Prof. Olcay Taner Yıldız

Funding: TUBITAK 1001

- Coordinating the project team, managing tasks, and tracking research progress.
- Performing linguistic annotations including word sense, morphological disambiguation, dependency parsing, argument detection (PropBank), frame detection (FrameNet), and semantic role labeling.
- Developing a consistent Abstract Meaning Representation (AMR) framework tailored for Turkish, aligned with project objectives.
- Creating a rule-based and an LLM-based parser for Turkish AMR graphs.

TECHNICAL & LANGUAGE SKILLS

Languages: Turkish (native), English (C1 – advanced), French (C1 – advanced), Spanish (A2 – beginner)

Programming Languages: Python (used for data analysis, NLP tasks, machine learning, deep learning), Java (used for OOP, data structures, designing and updating rule-based algorithms in NLP-Toolkit); SQL basics

Natural Language Processing: Experience with linguistic annotation, corpora management, and NLP libraries including SpaCy, NLP-Toolkit; morphological analysis, word sense disambiguation, semantic role labeling, dependency parsing, Abstract Meaning Representation (AMR)

Machine Learning / AI: PyTorch, scikit-learn; designing and training models, inner workings of Large Language Models

Data Analysis & Visualization: Pandas, NumPy, Matplotlib, Seaborn

Version Control & Collaboration: Git, GitHub

Development Tools: Jupyter Notebook/Lab, Google Colab, PyCharm, IntelliJ IDEA

PUBLICATIONS

Manuscripts in Preparation

- Cesur, Neslihan et al. (2025). *A unified Turkish treebank resource: morphosyntactic feature integration across multiple corpora*. (under review)
- Cesur, Neslihan et al. (2026). *Semantic Frame Classification and Semantic Role Labeling with Large Language Models*. (in progress)
- Cesur, Neslihan et al. (2026). *Developing Turkish AMR: Evaluating Rule-Based and LLM-Based AMR Parsing Approaches*. (in progress)

Conference Papers

- Cesur, Neslihan et al. (2026). *Evaluating Dependency Parse Quality Across Word-Level and Morpheme-Level Tokenizations*. (in progress)
- Cesur, Neslihan et al. (2025). *A Semi-Automated Approach to the Annotation of Argument Structures in Turkish Datasets*. In Proceedings of the 13th Global Wordnet Conference (GWC 2025). (to be published)

- Cesur, Neslihan et al. (2024). *Building Annotated Parallel Corpora Using the ATIS Dataset: Two UD-style Treebanks in English and Turkish*. In Proceedings of the 17th Workshop on Building and Using Comparable Corpora (LREC-COLING 2024), pp. 104–110.
- Arıcan, Bilge et al. (2022). *Morpholex Turkish: A Morphological Lexicon for Turkish*. In Proceedings of Globalex Workshop on Linked Lexicography, LREC 2022, pp. 68–74.
- Yenice, Arife B. et al. (2022). *Introducing StarDust: A UD-based Dependency Annotation Tool*. In Proceedings of the 16th Linguistic Annotation Workshop (LAW-XVI), LREC 2022, pp. 79–84.
- Marşan, Büşra, Oğuz K. Yıldız, et al. (2022). *A Learning-Based Dependency to Constituency Conversion Algorithm for the Turkish Language*. In Proceedings of LREC 2022, pp. 5054–5062.
- Marşan, Büşra, Neslihan Kara, et al. (2021). *Building the Turkish FrameNet*. In Proceedings of the 11th Global Wordnet Conference, pp. 118–125.
- Kuzgun, Aslı, Oğuz Kerem Yıldız, et al. (2021). *From Constituency to UD-Style Dependency: Building the First Conversion Tool for Turkish*. In Proceedings of RANLP 2021, pp. 761–769.
- Kara, Neslihan et al. (2020). *Creating a Syntactically Felicitous Constituency Treebank for Turkish*. In 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), IEEE, pp. 1–6.
- Kuzgun, Aslı, Neslihan Cesur, et al. (2020). *On Building the Largest and Cross-Linguistic Turkish Dependency Corpus*. In 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), IEEE, pp. 1–6.