

Soroosh Akef

PH.D. STUDENT · LINGUISTICS

Center of Linguistics of the University of Lisbon

+351 93-758-2379 | ✉ sorooshakef@edu.ulisboa.pt | 🏠 <https://www.clul.ulisboa.pt/en/pessoa/sorooshakef>

in [akef](#)

Education

University of Lisbon

Lisbon, Portugal

PH.D. IN LINGUISTICS

September 2022 - Present

- Dissertation Title: Fostering Portuguese Heritage Language Acquisition Through Proficiency-Tailored Input
- Supervisors: Prof. Dr. Amália Mendes, Prof. Dr. Detmar Meurers, Prof. Dr. Patrick Rebuschat

Sharif University of Technology

Tehran, Iran

M.SC. IN COMPUTATIONAL LINGUISTICS

September 2018 - February 2021

- Thesis Title: Automatic Difficulty Estimation of Thematic Similarity Multiple-Choice Questions
- Supervisors: Prof. Dr. Hossein Sameti, Prof. Dr. Mohammad Hadi Bokaei
- GPA: 15.95/20

Ferdowsi University of Mashhad

Mashhad, Iran

B.A. IN ENGLISH LANGUAGE AND LITERATURE

September 2014 - July 2018

- GPA: 18.82/20

Professional Experience

Natural Language Processing Intern

Remote

ASR GOOYESH PARDAZ

November 2020 - January 2021

- *Tasks:* Scraping Persian articles and their tags from the web and performing multi-label classification of article tags using a variety of traditional machine learning and deep learning approaches

English Language Teacher

Mashhad, Iran

KISH NOVIN LANGUAGE INSTITUTE

June 2018 - August 2018

- *Tasks:* Teaching two classes of intermediate-level adolescents

English Language Teacher

Mashhad, Iran

SHOKOOH ENGLISH INSTITUTE

June 2012 - September 2016

- *Tasks:* Teaching a range of classes from lower-intermediate- to advanced-level adolescents and adults

Publications

JOURNAL ARTICLES

Akef, S., Bokaei, M., & Sameti, H. (2020). Thematic similarity multiple-choice question answering with Doc2Vec: A step toward metaphorical language processing. *International Journal of Information & Communication Technology Research*, 12 (2), 46-53. <https://ijict.itrc.ac.ir/article-1-459-en.html>

INTERNATIONAL CONFERENCE PROCEEDINGS (PEER-REVIEWED)

Akef, S., Meurers, D., Mendes, A., & Rebuschat, P. (2025, March). Interpretable machine learning for societal language identification: Modeling English and German influences on Portuguese heritage language. In R. Muñoz Sánchez, D. Alfter, E. Volodina, & J. Kallas (Eds.), *Proceedings of the 14th Workshop on Natural Language Processing for Computer Assisted Language Learning* (pp. 50–62). University of Tartu Library. <https://aclanthology.org/2025.nlp4call-1.4/>

Akef, S., Mendes, A., Meurers, D., & Rebuschat, P. (2024, March). Investigating the generalizability of Portuguese readability assessment models trained using linguistic complexity features. In P. Gamallo, D. Claro, A. Teixeira, L. Real, M. Garcia, H. Gonçalves Oliveira, & R. Amaro (Eds.), *Proceedings of the 16th International Conference on Computational Processing of Portuguese* (pp. 332-341). Association for Computational Linguistics. <https://aclanthology.org/2024.propor-1.34>

- Akef, S.** & Bahrani, M. (2021) Fine-Tuning BERT for Persian poet identification. In *20th International Conference on Recent Research in Science and Technology*. Civilica. <https://civilica.com/doc/1412278/>
- Akef, S.**, Bokaei, M. H., & Sameti, H. (2020). Training Doc2Vec on a corpus of Persian poems to answer thematic similarity multiple-choice questions. In *10th International Symposium on Telecommunications (IST)* (pp. 146-149). IEEE. <https://doi.org/10.1109/IST50524.2020.9345918>
- Akef, S.**, & Bokaei, M. H. (2020). Answering poetic verses' thematic similarity multiple-choice questions with BERT. In *28th Iranian Conference on Electrical Engineering (ICEE)* (pp. 1-4). IEEE. <https://doi.org/10.1109/ICEE50131.2020.9260736>

Honors and Awards

- 2025 **Finalist of the ULisboa 3-Minute Thesis competition 2025**, University of Lisbon
Winner of the April 2025 LEAD Retreat Science Slam competition, LEAD Graduate School and Research Network
- 2022 **Ph.D. scholarship for the project “Promoção da Aquisição e Ensino do Português como Língua de Herança através de Ferramentas Digitais Inteligentes”**, Instituto Camões
- 2019 **Qualified for student financial aid**, Iran's National Elites Foundation
- 2018 **Rank 1 of the Iranian national master's degree entrance exam of linguistics**, National Organization of Educational Testing (Sanjesh)
Rank 9 of the Iranian national master's degree entrance exam of translation studies, National Organization of Educational Testing (Sanjesh)
Entrance exam waiver for admission to the linguistics master's degree program, Ferdowsi University of Mashhad

Presentations

CONFERENCE PRESENTATIONS

- Akef, S.**, Meurers, D., Mendes, A., & Rebuschat, P. (2025, June). *Investigating the Influence of English and German as Societal Languages on Portuguese as a Heritage Language Through Interpretable Machine Learning*. [Conference presentation]. The 34th Conference of the European Second Language Association (EuroSLA 34), Tromsø, Norway. https://en.uit.no/tavla/artikkel/845263/eurosla_34
- Akef, S.**, Meurers, D., Mendes, A., & Rebuschat, P. (2025, March). *Interpretable Machine Learning for Societal Language Identification: Modeling English and German Influences on Portuguese Heritage Language*. [Paper presentation]. 14th Workshop on Natural Language Processing for Computer Assisted Language Learning (NLP4CALL), Tallinn, Estonia. <https://spraakbanken.gu.se/en/research/themes/icall/nlp4call-workshop-series/nlp4call2025>
- Akef, S.**, Meurers, D., Mendes, A., & Rebuschat, P. (2024, October). *Studying the Role of L1 in Characterizing Portuguese L2 Proficiency Using Machine Learning*. [Conference presentation]. 40.º Encontro Nacional da Associação Portuguesa de Linguística, Ponta Delgada, Portugal. <https://apl.pt/eventos-realizados/40enapl/>
- Akef, S.**, Meurers, D., Mendes, A., & Rebuschat, P. (2024, September). *Empirically testing the “Trade-Off Hypothesis”: A machine learning experiment*. [Conference presentation]. Learner Corpus Research (LCR) 2024, Tartu, Estonia. <https://lcr2024.ut.ee/>
- Akef, S.**, Meurers, D., Mendes, A., & Rebuschat, P. (2024, July). *Investigating the contribution of accuracy and complexity in characterizing language proficiency: a machine learning experiment*. [Conference presentation]. The 33rd Conference of the European Second Language Association (EuroSLA 33), Montpellier, France. <https://eurosla33.sciencesconf.org/>
- Akef, S.**, Mendes, A., Meurers, D., & Rebuschat, P. (2024, March). *Investigating the generalizability of Portuguese readability assessment models trained using linguistic complexity features*. [Paper presentation]. 16th International Conference on Computational Processing of Portuguese, Santiago de Compostela, Galicia, Spain. <https://aclanthology.org/2024.propor-1.34>

- Akef, S.,** Mendes, A., Meurers, D., & Rebuschat, P. (2023, October). *Linguistic complexity features for automatic Portuguese readability assessment*. [Conference presentation]. XXXIX Encontro Nacional da Associação Portuguesa de Linguística, Covilhã, Portugal. <https://apl.pt/xxxix-enapl-2023/>
- Öksüz, D., **Akef, S.,** Mendes, A., & Rebuschat, P. (2023, September). *Tracking the development of noun-adjective collocations in L2 learners of Portuguese: A learner corpus analysis*. [Conference presentation]. The 32nd Conference of the European Second Language Association (EuroSLA 32), Birmingham, U.K. <https://www.birmingham.ac.uk/research/lacab/events/eurosla-32>
- Akef, S.,** Mendes, A., Meurers, D., & Rebuschat, P. (2023, July). *Automatic readability assessment of Portuguese: A linguistic complexity informed approach*. [Conference presentation]. L2 Portuguese @ Lancaster (L2P@L), Lancaster, U.K. <https://www.lancaster.ac.uk/heritage-language/l2pl2023/>
- Akef, S.** & Bahrani, M. (2021). *Fine-Tuning BERT for Persian poet identification*. [Paper presentation]. 20th International Conference on Recent Research in Science and Technology, Kerman, Iran. <https://civilica.com/doc/1412278/>
- Akef, S.,** Bokaei, M. H., & Sameti, H. (2020). *Training Doc2Vec on a corpus of Persian poems to answer thematic similarity multiple-choice questions*. [Paper presentation]. 10th International Symposium on Telecommunications (IST), Tehran, Iran. <https://doi.org/10.1109/IST50524.2020.9345918>
- Akef, S.,** & Bokaei, M. H. (2020). *Answering poetic verses' thematic similarity multiple-choice questions with BERT*. [Paper presentation]. 28th Iranian Conference on Electrical Engineering (ICEE), Tabriz, Iran. <https://doi.org/10.1109/ICEE50131.2020.9260736>

Skills

TECHNICAL SKILLS

Programming: Python, Javascript, R, Java
 Web Development: HTML, CSS
 Natural Language Processing
 Machine Learning
 Deep Learning
 Web Scraping
 Large Language Model Fine-Tuning
 Version Control
 Exploratory Data Analysis
 Bayesian Statistics
 Inferential Statistics

LANGUAGES

English (C2)
 Persian (L1)
 Portuguese (B1)
 Spanish (A2)
 German (A1)