**Mohammad Sadegh Rasooli**
Postdoctoral Researcher
University of Pennsylvania
rasooli@seas.upenn.edu

## RESEARCH INTERESTS

Natural Language Processing, Machine Translation, Deep Learning for Natural Language Processing.

## EMPLOYMENT

- Postdoctoral Researcher, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA, 2020-Now. Supervisor: Chris Callison-Burch.
- Research scientist, Facebook AI, Menlo Park, CA, 2018-2020.
- Research assistant, Department of Computer Science, Columbia University, 2012-2018.
- Research intern, Microsoft Artificial Intelligence and Research, Sunnyvale, CA, Summer 2017.
- Research intern, Google Research, New York, NY, Summer 2015.
- Research intern, Yahoo Research, New York, NY, Summer 2014.
- Research intern, Nuance Communications, Sunnyvale, CA, Summer 2013.
- Research manager, Dadegan Research Group, Tehran, Iran, 2010-2012.
- Research engineer, Noor, Tehran, Iran, 2009-2010.

## EDUCATION

- Ph.D., Computer Science, Department of Computer Science, Columbia University, 2012-2019.
  - Advisor: Michael Collins.
  - *Thesis*: Cross-Lingual Transfer of Natural Language Processing Systems
- MSc. (Ph.D. track), Department of Computer Science, Columbia University, 2012-2014.
- MSc., Artificial Intelligence, Iran University of Science and Technology, 2009-2012.
- BSc., Software Engineering, Iran University of Science and Technology, 2005-2009.

## PUBLICATIONS

Google Scholar page: `http://scholar.google.com/citations?user=UVTYgJIAAAAJ`

1. **Mohammad Sadegh Rasooli**, Pegah Safari, Amirsaeid Moloodi, and Alireza Nourian. The Persian Dependency Treebank Made Universal. arXiv:2009.10205, Sep. 2020.
2. Maryam Aminian, **Mohammad Sadegh Rasooli**, and Mona Diab. Multitask Learning for Cross-Lingual Transfer of Broad-coverage Semantic Dependencies. EMNLP 2020.
3. **Mohammad Sadegh Rasooli**, and Michael Collins. Low-Resource Syntactic Transfer with Unsupervised Source Reordering, NAACL 2019.
4. Maryam Aminian, **Mohammad Sadegh Rasooli**, and Mona Diab. Cross-Lingual Transfer of Semantic Roles: From Raw Text to Semantic Roles. The 13th International Conference on Computational Semantics (IWCS 2019), Gothenburg, Sweden, May 2019.
5. **Mohammad Sadegh Rasooli**, Cross-Lingual Transfer of Natural Language Processing Systems, PhD Thesis, Columbia University, 2019.
6. **Mohammad Sadegh Rasooli**, and Sarangarajan Parthasarathy.Entity-Aware Language Model as an Unsupervised Reranker, INTERSPEECH, Heyderabad, India, Spetember 2018. Link: `https://arxiv.org/abs/1803.04291`
7. **Mohammad Sadegh Rasooli**, Noura Farra, Axinia Radeva, Tao Yu, and Kathleen McKeown. Cross-lingual sentiment transfer with limited resources. Machine Translation, issn=1573-0573, doi=10.1007/s10590-017-9202-6, November 2017. `https://link.springer.com/article/10.1007/s10590-017-9202-6`.

8. Maryam Aminian, **Mohammad Sadegh Rasooli** and Mona Diab. Transferring Semantic Roles Using Translation and Syntactic Information. Proceedings of the The 8th International Joint Conference on Natural Language Processing, pages 1319, Taipei, Taiwan, November 27 December 1, 2017. http://aclweb.org/anthology/I17-2003.
9. **Mohammad Sadegh Rasooli** and Michael Collins. Cross-Lingual Syntactic Transfer with Limited Resources. Transactions for the association for computational linguistics, vol. 5, pp. 279293, 2017. http://www.aclweb.org/anthology/Q17-1020.
10. **Mohammad Sadegh Rasooli** and Michael Collins. Density-Driven Cross-Lingual Transfer of Dependency Parsers. Conference on Empirical Methods in Natural Language Processing (EMNLP), Lisboa, Portugal, September 2015. Link: http://www.aclweb.org/old_anthology/D/D15/D15-1039.pdf
11. Alireza Nourian, **Mohammad Sadegh Rasooli**, Mohsen Imany and Heshaam Faili. On the Importance of Ezafe Construction in Persian Parsing. The 53rd Annual Meeting of the Association for Computational Linguistics (ACL) and the 7h International Joint Conference on Natural Language Processing (IJCNLP), Beijing, China, July 2015. Link: http://anthology.aclweb.org/P/P15/P15-2144.pdf
12. **Mohammad Sadegh Rasooli** and Joel Tetreault. Yara Parser: A Fast and Accurate Dependency Parser, 23 Mar 2015. Link: arXiv:1503.06733v1
13. **Mohammad Sadegh Rasooli**, Manouchehr Kouhestani, and Amirsaeid Moloodi. Persian Syntactic Treebank: a Research Based on Dependency Grammar. Supreme Council of Information and Communication Technology (SCICT), Tehran, Iran, 2014 (in Persian, ISBN: 978-964-8846-37-9). Link: http://dadegan.ir/catalog/perdt
14. Xiaodong Cui, Brian Kingsbury, Jia Cui, Bhuvana Ramabhadran, Andrew Rosenberg, **Mohammad Sadegh Rasooli**, Owen Rambow, Nizar Habash and Vaibhava Goel. Improving Deep Neural Network Acoustic Modeling For Audio Corpus Indexing Under The IARPA Babel Program. 15th Annual Conference of the International Speech Communication Association (INTERSPEECH), pp. 2103–2107, Singapore, Septemeber 2014. Link: http://193.6.4.39/~czap/letoltes/IS14/IS2014/PDF/AUTHOR/IS141114.PDF
15. **Mohammad Sadegh Rasooli**, Thomas Lippincott, Nizar Habash and Owen Rambow. Unsupervised Morphology-Based Vocabulary Expansion. The 52nd Annual Meeting of the Association for Computational Linguistics (ACL), pp. 1349–1359, Baltimore, Maryland, USA, June 2014. Link: http://www.anthology.aclweb.org/P/P14/P14-1127.pdf
16. **Mohammad Sadegh Rasooli** and Joel Tetreault. Non-Monotonic Parsing of Fluent Umm I mean Disfluent Sentences. 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL), pp. 48–53, Gutenberg, Sweden, April 2014. Link: http://anthology.aclweb.org/E/E14/E14-4010.pdf
17. **Mohammad Sadegh Rasooli** and Joel Tetreault. Joint Parsing and Disfluency Detection in Linear Time, Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 124-129, Seattle, Washington, October 2013. Link: http://www.aclweb.org/old_anthology/D/D13/D13-1013.pdf
18. **Mohammad Sadegh Rasooli**, Ahmed El Kholy and and Nizar Habash. Orthographic and Morphological Processing for Persian to English Statistical Machine Translation, The 6th International Joint Conference on Natural Language Processing (IJCNLP), pp. 1047-1051, Nagoya, Japan, October 2013. Link: http://www.anthology.aclweb.org/I/I13/I13-1144.pdf
19. **Mohammad Sadegh Rasooli**, Manouchehr Kouhestani, and Amirsaeid Moloodi. Development of a Persian Syntactic Dependency Treebank, The 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL–HLT), pp. 306-314, Atlanta, Georgia, USA, June 2013. Link: https://pdfs.semanticscholar.org/4719/86880bdc1ce94e6e8120cb81403c2b3fca18.pdf
20. Maryam Aminian, **Mohammad Sadegh Rasooli** and Hossein Sameti. Unsupervised Induction of Persian Semantic Verb Classes Based on Syntactic Information, International Conference on Language Processing and Intelligent Information Systems (The 20th International Conference on Intelligent Information Systems), pp. 112-124, Warsaw, Poland, June 2013. Link: http://link.springer.com/chapter/10.1007/978-3-642-38634-3_13

21. **Mohammad Sadegh Rasooli**, Heshaam Faili, Behrouz Minaei-Bidgoli and Maryam Aminian. Unsupervised Extraction of Verb Valency in Persian, Journal of Signal and Data Processing, 2(18), pp. 3-12, 2013 (in Persian). Link: `http://jsdp.rcisp.ac.ir/files/site1/user_files_60a4f6/rasoolims-A-10-302-1-bf3a626.pdf`
22. **Mohammad Sadegh Rasooli** and Heshaam Faili. Fast Unsupervised Dependency Parsing with Arc-Standard Transitions, in EACL workshop on ROBUS-UNSUP 2012: Joint Workshop on Unsupervised and Semi-Supervised Learning in NLP, pp. 1-9, Avignon, France, 2012. Link: `http://anthology.aclweb.org/W/W12/W12-0701.pdf`
23. Manouchehr Kouhestani, Amirsaeid Moloodi, and **Mohammad Sadegh Rasooli**. Persian Verb Valency Lexicon: An Attempt Toward Teaching Persian to Non-native Persian Speakers, International Conference on Spread of Persian Language and Literature, pp. 913-930, Tehran, Iran, 2012 (In Persian). Link: `http://dadegan.ir/sites/default/files/articles/Persian-1390-Persian_Valency_Lexicon.pdf`
24. **Mohammad Sadegh Rasooli**, Amirsaeid Moloodi, Manouchehr Kouhestani, and Behrouz Minaei-Bidgoli. A Syntactic Valency Lexicon for Persian Verbs: The First Steps towards Persian Dependency Treebank, in 5th Language & Technology Conference (LTC): Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 227-231, Poznań, Poland, 2011. Link: `http://www.cs.columbia.edu/~rasooli/papers/VAL-1.pdf`
25. **Mohammad Sadegh Rasooli**, Heshaam Faili, and Behrouz Minaei-Bidgoli. Unsupervised Identification of Persian Compound Verbs. in 10th Mexican International Conference on Artificial Intelligence (MICAI), pp. 394-406, Puebla, Mexico, 2011. Link: `http://link.springer.com/chapter/10.1007%2F978-3-642-25324-9_34`
26. **Mohammad Sadegh Rasooli**, Omid Kashefi, and Behrouz Minaei-Bidgoli. Extracting Parallel Paragraphs and Sentences from English-Persian Translated Documents. in The Seventh Asia Information Retrieval Societies Conference (AIRS 2011), pp. 574-583, University of Wollongong in Dubai, Dubai, United Arab Emirates, 2011. Link: `http://link.springer.com/chapter/10.1007%2F978-3-642-25631-8_52`
27. **Mohammad Sadegh Rasooli**, Omid Kashefi, and Behrouz Minaei-Bidgoli. Effect of Adaptive Spell Checking in Persian, in 7th IEEE International Conference on Natural Language Processing and Knowledge Engineering (NLPKE'11), pp. 161-164, Tokushima, Japan, 2011. Link: `https://www.researchgate.net/profile/Omid_Kashefi/publication/220875315_Effect_of_adaptive_spell_checking_in_Persian/links/0deec518f60ab2f20f000000.pdf`
28. **Mohammad Sadegh Rasooli** and Behrouz Minaei-Bidgoli. A New Approach for Persian Spell Checking, in the 2nd Data Mining Conference (IDMC08), Tehran, Iran, 2008 (in Persian). Link: `http://www.cs.columbia.edu/~rasooli/papers/AnewapproachforPersianspellchecking.pdf`

## HONORS

- Ranked %0.6 among about 400,000 competitors of national B.Sc. university entrance exam in Math and Physics, 2005.
- Ranked %0.36 among about 250,000 competitors of national B.Sc. university entrance exam in English language, 2005.

## PROFESSIONAL SERVICES

**ACL-related reviewing:** EMNLP 2015, NAACL 2016, ACL 2016, EMNLP 2016, ACL 2017, EMNLP 2017, NAACL 2018, EMNLP 2018, ACL 2018, NAACL 2019, ACL 2019, EMNLP 2019, CoNLL 2019, ACL 2020, EMNLP 2020, AACL 2020, CONLL 2020, Computational Linguistics, ACM Transactions on Asian Language Information Processing (TALIP), Language Resources and Evaluation (LRE), Transactions of the Association for Computational Linguistics (TACL), Elsevier Computer Speech and Language.

**Conference reviewing:** Symposium on Artificial Intelligence and Signal Processing (AISP 2013), 19th International Conferences of Computer Society of Iran (CSICC 2013), 1st National Conference on the Semantics ans Syntax of Iranian Linguistics 2014.

## TECHNICAL SKILLS

**Programming Languages:** Java, Python, C#, C++.

## TEACHING EXPERIENCE

- Co-instructor, natural language processing, Department of Computer Science, Columbia University, Fall 2017.
  - Teacher: Michael Collins
- TA, natural language processing, Department of Computer Science, Columbia University, Spring 2015.
  - Teacher: Michael Collins
- TA, natural language processing, Department of Computer Science, Columbia University, Fall 2013.
  - Teacher: Michael Collins
- TA, Artificial Intelligence, Iran University of Science and Technology, 2009-2011.
  - Teacher: Behrouz Minaei-Bidgoli

## LANGUAGES

**Persian:** Native; **English**: Fluent; **Arabic**: Grammar and reading fair.