

# Marzena Karpinska

## Curriculum Vitae

☎ (+1) 413-316-4261  
✉ [marzena.karpinska@gmail.com](mailto:marzena.karpinska@gmail.com)  
in [marzena-karpinska-93049487](#)  
G [Google Scholar](#)

Marzena Karpinska received her Ph.D. in linguistics from the Department of Language and Information Sciences, the University of Tokyo. Her research interests lie at the intersection of natural language processing, theoretical linguistics, and language acquisition.

### Education

- 2014-2020 Ph.D.: The University of Tokyo (Japan), Language and Information Sciences Dept.  
Thesis *The influence of socio-indexical cues on speech perception by non-native listeners.*
- 2012-2014 M.A.: The University of Tokyo (Japan), Language and Information Sciences Dept.  
Thesis *The Source-Parth-Goal Schema in Japanese and Polish.*
- 2007-2010 M.A.: The University of Warsaw (Poland), Oriental Studies Dept., Japanese Linguistics  
Thesis *Semantics of Japanese Evaluative Adjectives.*
- 2004-2007 B.A.: The University of Warsaw (Poland), Oriental Studies Dept., Japanese Linguistics  
Thesis *Semantics of Japanese Transitive - Intransitive Verb Pair 'ateru:ataru.'*

### Research Stays

- 2011-2012 The University of Tokyo (Japan), Language and Information Sciences Dept.  
Focus: *The Conceptualization of 'Source' in Japanese and Polish.*
- 2008-2009 Tokyo University of Foreign Studies (Japan), Research Student of Japanese Language  
Focus: *Coexisting Intransitive Verbs: A Case of Japanese.*

### Professional Experience

- 2018–2020 Research Assistant: The University of Tokyo, Japan (part-time position), A joint research project of The University of Tokyo, Tokyo Institute of Technology, and Konan Univeristy  
Tasks: Cleaning and lemmatizing multiple corpora  
Training word2vec models and aligning embeddings with MUSE  
Creating cognates dataset for Romance, Germanic, and Slavic languages
- 2016-2017 Research Assistant: The University of Tokyo, Japan (part-time position)  
Tasks: Preparing auditory stimuli, designing acoustic perception experiments in SuperLab  
Collecting participant data, statistical analysis in R
- 2014– Language Specialist: Innovative Language Learning (ILL), Japan (part-time position)  
Tasks: Developing curricula, assessments, study materials for over 30 languages on the ILL digital learning platform ([URL](#))  
Designing experiments to measure the students' progress and analyzing trends in their performance  
Improving user experience with the Alexa voice assistant for the ILL language skill
- 2014-2018 Instructor: Nippon Institute of Technology, Japan (part-time position)  
Courses: English for System Engineering, English for Programming, English for Design
- 2008– Japanese-Polish-English Translator and Interpreter (freelance)

Experience Various translations including manga, legal documents, technical documentation and oral presentations for many high profile clients such as Deloitte, Sony, Toshiba, Toyota and Polish Embassy in Japan

## Awards and scholarships

- Research Grant for Doctoral Candidates, 2017-2019 (granted yearly)
- Research Student Scholarship of Ministry of Education of Japan, 2008-2009 and 2011-2017
- Scholarship awarded yearly by Rector of The University of Warsaw, 2005-2010

## Skills and Expertise

- *Programming*: Python, R, JavaScript;
- *Machine learning frameworks*: scikit-learn, PyTorch;
- *Theoretical frameworks*: distributional semantics, frame semantics, sociolinguistics, pragmatics, cognitive linguistics, discourse analysis, sociophonetics, phonetics, cognitive semantics;
- *Research Skills*: designing controlled experiments, collecting participant data, statistical analysis, ethics of data collection;
- *Languages*: English, Polish, Japanese (native-level), German (intermediate), Chinese, Czech, Russian, French (beginner), broad knowledge of typological traits of over 30 languages available on the ILL platform ([URL](#))

## Publications

- Marzena Karpinska, Nader Akoury, and Mohit Iyyer (2021). *The Perils of Using Mechanical Turk to Evaluate Open-Ended Text Generation*. arXiv: 2109.06835 [cs.CL]
- Marzena Karpinska (2019). “How accented do Caucasian-looking vs. Asian-looking native speakers sound to a Japanese listener?” *Proceedings of the The International Congress of Phonetic Sciences*. Melbourne, Australia: ICPhS, pp. 3691–3695. URL: [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS\\_3740.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS_3740.pdf)
- Marzena Karpinska, Paula Kurzawska, and Katarzyna Rozanska (2019). “Digital Lingua Franca or a Culture-Specific Product Leading to Misunderstandings?” *Emoticons, Kaomoji, and Emoji: The Transformation of Communication in the Digital Age (Routledge Research in Language and Communication)*. Ed. by E. GiannoulisLukas and R.A. Wilde. Routledge. Chap. 4
- Marzena Karpinska, Bofang Li, Anna Rogers, and Aleksandr Drozd (2018). “Subcharacter Information in Japanese Embeddings: When Is It Worth It?” *Proceedings of the Workshop on the Relevance of Linguistic Structure in Neural Architectures for NLP*. Melbourne, Australia: ACL, pp. 28–37. URL: <http://aclweb.org/anthology/W18-2905>
- Marzena Karpinska, Shodai Uchida, and Isabelle Grenon (2015). “Vowel perception by listeners from different English dialects”. *Proceedings of the The International Congress of Phonetic Sciences*. Glasgow, Scotland: ICPhS. URL: <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0264.pdf>