

Hadas KOTEK

hkotek@alum.mit.edu

<http://hkotek.com>

(203) 415-0258

SUMMARY

Language scientist with 10+ years experience in experimental and theoretical linguistics, specializing in natural language generation, understanding, and processing. Expertise in experiment design, language data collection and analysis, data annotation, taxonomy design, and team and budget management. Outstanding verbal, analytical, leadership, and teamwork skills.

SKILLS

Data science Taxonomy and ontology design | Annotation guideline creation | Experiment design | Data collection, annotation, processing, visualization, statistical analysis | Qualitative and quantitative research methods

Project management Team work and team leadership | Personnel, budget management | Grant writing | Written, oral presentation | Problem solving

NLP knowledge Named Entity Recognition | Part of Speech Tagging | Constituency Parsing | Dependency Parsing | Discourse-Sensitive Parsing, multi-turn conversations

Technical skills Python | R | SPSS | Markdown | SQL (basic) | HTML (basic) | Javascript (basic) | Pascal | \LaTeX | Advanced Excel | Praat

Languages Hebrew (native) | English (native-like) | German (fluent, ZOP certificate) | Arabic (conversational) | Spanish (conversational) | Japanese (beginner) | French (beginner)

EDUCATION

2014 *PhD in Linguistics, Massachusetts Institute of Technology, Cambridge, MA.*

2007 *BA in Linguistics and Political Science, Tel-Aviv University, Israel, summa cum laude.*

Computer Science courses taken: Infinitesimal Calculus 1+2, Discrete Mathematics, Computer Science, Math, Physics, and Language AP courses taken in high school.

WORK EXPERIENCE

Siri – Natural Language Annotation Project Lead,

Natural Language Understanding (Data Science team) Apple, Inc

2019 – present

- Leader of annotation teams in several locales. Developed metrics to measure annotator performance (accuracy, throughput), providing feedback and guidance for improvement.
- Facilitator of cross-functional collaboration with data scientists and engineers, contributing to R&D projects, design and deployment of all Siri Understanding annotation projects.
- Prioritized competing projects; set and maintained a schedule for milestone and project completion to align with organizational goals and timelines.
- Led ontology design and development to contribute to existing products, as well as products in development. Created annotation guidelines and trainings for Siri and Siri International.
- Coordinated with vendors and Apple Finance to establish long-term partnerships and ensure proper and accurate invoicing.

Faculty member in Linguistics 2014 – 2019
Yale University (2018 – 2019, 2016 – 2017), **New York University** (2017 – 2018)
McGill University (2014 – 2016)

- Collected linguistic data and developed cutting-edge linguistic theories, resulting in:
 - 1 book ([MIT Press](#)); 1 edited volume; over 20 papers and 60 presentations in international conferences and workshops; 10 grants and awards (including prestigious Mellon Fellowship, \$121,000).
- Won grant to lead 25-person team in large-scale data collection project.
 - Designed study, created and managed budget, and hired team; created [annotation guidelines](#) and supervised annotators, resulting in corpus of 16K sentences from four leading journals; implemented analysis script in R to study gender representation in Linguistics over 20 years, resulting in a paper and position paper.
- Co-organized 6 international conferences and workshops with hundreds of attendees.
- Created teaching materials; taught lectures, advanced courses and seminars; managed Teaching Assistant teams (14 courses in syntax, semantics, and experimental linguistics, with a total of over 500 enrolled students.)
- Advised PhD dissertations, graduate Qualifying Papers, and undergraduate theses.

Graduate student researcher, Massachusetts Institute of Technology 2010 – 2014

- Synthesized a large amount of natural language data from several languages, including English, German, and Japanese. Wrote a thesis on the structure and meaning of questions.
- Designed and implemented psycholinguistics, neurolinguistic, sentence processing, first language acquisition, and large-scale grammaticality judgment experiments. Created data visualization and analysis scripts (samples [design document 1](#), [design document 2](#)).
- 15 papers, 37 presentations, 10 grants and awards, including NSF Grant.

Co-creator of *Turktools*, Massachusetts Institute of Technology 2012 – 2013

- Free toolkit for implementation and analysis of experiments on Amazon Mechanical Turk.
 - Designed and coded HTML scripts and R script for data visualization and analysis.
 - Designed and contributed to programming of Python scripts.
 - 1 peer-reviewed paper in top journal, 1 grant, 1 workshop at MIT.

Machine Learning Coordinator, Ontology Ltd, Tel-Aviv, Israel 2005 – 2006

- Built a knowledge graph for machine learning software based on a multilingual classification algorithm. Platform targeted cross-language/culture economic relations in the EU.
- Contributed to taxonomy design.
- Collected, processed, and annotated datasets for the Spanish component of the software.

Translator, Israeli Prime Minister's Office and Israeli Defense Forces 2001 – 2003

- Arabic-Hebrew translator: audio, texts, and video sources.
- Contributed to wartime-specialized automatic Machine Translation software.

Link to academic CV: <https://hkotek.com/KotekCV.pdf>

Links to sample scripts and design documents: <https://github.com/hkotek/scripts>