

Dorottya Demszky

Margaret Jacks Hall, Stanford, CA 94305-2150 ♠ ddemszky@stanford.edu

EDUCATION

Stanford University

September 2017 – Present

- PhD Candidate in Linguistics
- Melvin and Joan Lane Stanford Graduate Fellow

Princeton University

September 2013 – June 2017

- BA in Linguistics, Minor in Computer Science
- Summa Cum Laude

LSA Summer Linguistics Institute, University of Chicago

July 2015

- Fellowship Recipient

PROFESSIONAL EXPERIENCE

Google

Summer 2020

Research Intern

Seattle, WA

- Worked on a multi dialect data collection project and helped design methods to measure dialect density in the data automatically.

Google

Summer 2019

Research Intern

Mountain View, CA

- Collected data, conducted analyses and built models for emotion and sentiment analysis.

Center for the Study of Language and Information (CSLI)

Summer 2016

Research Intern

Stanford University

- Worked with Professor Percy Liang and graduate student Kelvin Guu on a project pertaining to understanding the linguistic properties of natural language inference.

PUBLICATIONS

- Demszky, D., Sharma, D., Clark, J. H., Prabhakaran, V., and Eisenstein, J. (2021). Learning to Recognize Dialect Features. In *Proceedings of the 19th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*
- Demszky, D. (2021). The Role of Verb Semantics in Hungarian Verb-Object Order. In *Proceedings of the Linguistic Society of America (LSA)*
- Lucy*, L., Demszky*, D., Bromley, P., and Jurafsky, D. (2020). Content Analysis of Textbooks via Natural Language Processing: Findings on Gender, Race, and Ethnicity in Texas US History Textbooks. *AERA Open*, 6(3):2332858420940312
- Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., and Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*
- Kmetty, Z., Vincze, V., Demszky, D., Ring, O., Nagy, B. N., and Szabó, M. K. (2020). *Pártélet*: A Hungarian Corpus of Propaganda Texts from the Hungarian Socialist Era. In *Proceedings of the 13th International Conference on Language Resources and Evaluation (LREC)*

- Demszky, D., Garg, N., Voigt, R., Zou, J., Gentzkow, M., Shapiro, J., and Jurafsky, D. (2019). Analyzing Polarization in Social Media: Method and Application to Tweets on 21 Mass Shootings. In *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*
- Demszky*, D., Guu*, K., and Liang, P. (2018). Transforming Question Answering Datasets Into Natural Language Inference Datasets. *arXiv preprint arXiv:1809.02922*

* indicates equal contribution

CONFERENCES & INVITED TALKS

NAACL, Mexico City (virtual)	June 2021
◦ Giving a talk on <i>Learning to Recognize Dialect Features</i> .	
Rising Stars in AI for Social Good Speaker Series, Harvard CRCS	April 2021
◦ Nominated to give a talk on my work.	
Cornell NLP Seminar	April 2021
◦ Gave a talk on my work on measuring uptake in student-teacher interactions.	
LSA Annual Meeting, San Francisco (virtual)	January 2021
◦ Gave a talk on <i>The Role of Verb Semantics in Hungarian Verb-Object Order</i> .	
Eötvös Loránd University, Hungary	December 2020
◦ Gave a talk on the <i>Content Analysis of Textbooks</i> .	
Conference on Educational Data Science, Stanford	September 2020
◦ Gave a talk with Lucy Li on <i>Content Analysis of Textbooks</i> .	
◦ Our submission received the Best Paper Award.	
10th Annual New Directions in Analyzing Text as Data (TADA)	October 2019
◦ Gave a talk with Lucy Li on <i>Content Analysis of Textbooks</i> .	
University of San Francisco Seminar Series in Data Science	October 2019
◦ Gave a talk on our paper <i>Analyzing Polarization in Social Media</i> .	
NAACL, Minneapolis	June 2019
◦ Gave a talk on <i>Analyzing Polarization in Social Media</i> .	
McGill Canadian Conference for Linguistics Undergraduates (McCCLU)	March 2013
◦ Gave a talk on Hungarian noun incorporation, based on a junior research paper advised by Professor Joshua Katz.	
Toronto Undergraduate Linguistics Conference (TULCON)	March 2013
◦ Gave a talk on Hungarian noun incorporation, based on a junior research paper advised by Professor Joshua Katz.	

TEACHING & MENTORING

Linguistics Department, Stanford University
Teaching Assistant

- I am the TA for Natural Language Understanding (AY Spring 2021), led by Prof. Chris Potts.

- I was the Corpus TA (AY 2019/20), as part of which I am managing corpora, holding workshops and helping students and faculty use corpora in their research.
- I was a TA for Introduction to Psycholinguistics (Autumn 2018), led by Prof. Judith Degen.

Petey Greene Prison Tutoring Program

January 2015 – May 2016

Volunteer Tutor & Course Co-Leader

- Tutored college-level Math and co-led an employment preparation course at correctional facilities in New Jersey.

Rockefeller College, Princeton University

September 2015 – May 2017

Peer Academic Advisor

- Advised a group of freshmen and sophomores on academic issues, such as adjustment to college, course selection and major choices.

ACTIVITIES

Tarisznya Táborok (Knapsack Camps)

2016 – Present

Co-Founder & Member of the Board of Trustees & Fundraising Chair

- Co-founded an NGO that seeks to empower underprivileged children in Hungary through inclusive educational summer camps and mentorship programs.
- Have raised more than \$50,000 since 2016 via grants and crowdfunding campaigns.

Hungarian Linguistics Olympiad

2012 – Present

Founder & Board Member

- Initiated and co-organized the first Hungarian Linguistics Olympiad in 2012.
- Have written and translated problems for the six-round contest, which has around 100 participants every year.
- Providing mentorship to participants.

North American Computational Linguistics Olympiad

2013 – 2017

Member of the Problem Committee and Organizing Committee

- Wrote and helped correct problems for the contest.
- Served as the site coordinator for Princeton University.

Horizons for Homeless Children

September 2012 – April 2013

Certified Playspace Activity Leader

Cambridge, MA

- Played and studied for school with homeless children at a homeless shelter for 3 hours/week.

ACADEMIC SERVICE

Association for Computational Linguistics

Program Committee Member & Reviewer

- Reviewed papers for ACL 2019, EMNLP 2020, ACL 2021, ACL SRW 2021.

American Education Research Association (AERA) Open Journal

Reviewer

- Reviewed for the data science special issue in 2020.

Linguistics Department, Stanford University

- Member of the Colloquium Committee

2019 – 2020

◦ Member of the Social Committee 2017 – 2018

Stanford NLP Group

September 2017 – Present

Social Co-Chair

- Co-organizing social events for the members of the Stanford NLP Group.

Princeton Linguistics Club

2015 – 2017

Co-President

- Co-organized themed meetings and discussions around linguistic topics for undergraduates.

SKILLS

Programming	Python, R, Java, Javascript, C
Languages	Hungarian, English, German, French, basic Russian

AWARDS & HONORS

Best Paper Award at the Conference on Education Data Science	2020
Melvin and Joan Lane Stanford Graduate Fellowship	2017 – 2022
Senior Thesis Prize (Princeton University)	2017
Phi Beta Kappa Honor Society	2017
Davis Projects for Peace Prize	2016
Dr. Elemér & Éva Kiss Scholarship	2013, 2014, 2015, 2016, 2017
Linguistics Summer Institute Fellowship by the Linguistics Society of America (LSA)	2015
Recipient of the Városmajori Medal for Academic Achievements (awarded to one or two members of the graduating class at Városmajori Gimnázium)	2012
Member of the Hungarian team at the 10th International Linguistics Olympiad (IOL)	2012
4th Place at the Hungarian National Linguistics Olympiad	2012
6th Place at the Hungarian National Philosophy Olympiad	2012
11th Place at the Hungarian National Philosophy Olympiad	2011
National finalist at the Jugend Debattiert International German Debating Competition	2012
3rd Prize at the National English Short Story Writing Contest organized by the American Embassy in Hungary	2012
Honorable Mention at the National Latin Translation Competition	2008