Tyler A. Chang

he/him/his
tachang@ucsd.edu / tylerachang.github.io

EDUCATION

University of California San Diego   2020-Present
PhD Student in Cognitive Science
Halıcıoğlu Data Science Institute, Graduate Fellow

Carleton College, Northfield, MN   2016-2020
BA in Mathematics, summa cum laude
BA in Cognitive Science, with distinction

Eötvös Loránd University, Budapest, Hungary   Fall 2018
Budapest Semesters in Mathematics

Doshisha University, Kyoto, Japan   Spring 2018
Carleton Linguistics and Culture Program

RESEARCH INTERESTS

- **Large language models**: pre-training dynamics, deep learning interpretability and analysis.
- **Cognitive linguistics**: distributional semantics, language acquisition.

INDUSTRY

Google Research, PhD Student Researcher   2022-Present
- Responsible AI: pre-training, tuning, and analysis of large language models.

Amazon Science, Applied Scientist Intern   2021, 2022
- AWS AI Labs: linguistic dataset drift and out-of-domain language model generalization.
- Alexa AI: content change prediction on the web for intelligent web crawling.

Google, Software Engineering Intern   2018, 2019, 2020
- Cloud AI Translation: automatic sentence pair extraction from human-translated documents.
- Geo Machine Perception (Google Maps, Earth): aerial and street level imagery alignment at scale.
- Chrome Web Store (Engineering Practicum Intern): building servers for tens of millions of users.

PUBLICATIONS   *Peer-reviewed.*

Note: top-tier venues in natural language processing are often conference proceedings rather than journals.

Chang, T. A., & Bergen, B. K. (2024). Language model behavior: A comprehensive survey. *Computational Linguistics*.

Unger, L., **Chang, T. A.**, Savic, O., Bergen, B. K., & Sloutsky, V. M. (2024). When is a word in good company for learning? *Developmental Science*.

Chang, T. A.*, Tomanek, K.*, Hoffmann, J., Thain, N., van Liemt, E., Meier-Hellstern, K., & Dixon, L. (2024). Detecting hallucination and coverage errors in retrieval augmented generation for controversial topics. *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources, and Evaluation* (LREC-COLING). *Equal contribution. Work done at Google Research.

Shah, C.$^\diamond$, Chandak, Y.$^\diamond$, Mane, A.$^\diamond$, Bergen, B. K., & **Chang, T. A.** (2024). Correlations between multilingual language model geometry and crosslingual transfer performance. *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources, and Evaluation* (LREC-COLING). $^\diamond$Undergraduate mentees.

Michaelov, J.*, Arnett, C.*, **Chang, T. A.**, & Bergen, B. K. (2023). Structural priming demonstrates abstract grammatical representations in multilingual language models. *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP). *Equal contribution.

Chang, T. A., Halder, K., Anna John, N., Vyas, Y., Benajiba, Y., Ballesteros, M., & Roth, D. (2023). Characterizing and measuring linguistic dataset drift. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics* (ACL). Work done at Amazon Science.

Trott, S.*, Jones, C. R.*, **Chang, T. A.**, Michaelov, J., & Bergen, B. K. (2023). Do large language models know what humans know? *Cognitive Science*. *Equal contribution.

Chang, T. A., Tu, Z., & Bergen, B. K. (2022). The geometry of multilingual language model representations. *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP).

Chang, T. A., & Bergen, B. K. (2022). Word acquisition in neural language models. *Transactions of the Association for Computational Linguistics* (TACL). Presented at ACL 2022.

Chang, T. A., & Bergen, B. K. (2022). Does contextual diversity hinder early word acquisition? *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (CogSci).

Jones, C. R., **Chang, T. A.**, Coulson, S., Michaelov, J., Trott, S., & Bergen, B. K. (2022). Distributional semantics still can't account for affordances. *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (CogSci).

Chang, T. A., Xu, Y., Xu, W., & Tu, Z. (2021). Convolutions and self-attention: Re-interpreting relative positions in pre-trained language models. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (ACL-IJCNLP).

Xu, Y., Xu, W., **Chang, T. A.**, & Tu, Z. (2021). Co-scale conv-attentional image transformers. *Proceedings of the IEEE/CVF International Conference on Computer Vision* (ICCV).

WORKSHOPS AND ABSTRACTS
<div align="right">Peer-reviewed.</div>

Arnett, C.*, **Chang, T. A.***, & Bergen, B. K. (2024). A bit of a problem: Measurement disparities in dataset sizes across languages. *Proceedings of the Annual Meeting of the Special Interest Group on Under-Resourced Languages* (workshop at LREC-COLING). *Equal contribution.

Arnett, C., **Chang, T. A.**, Michaelov, J., & Bergen, B. K. (2023). Crosslingual structural priming and the pre-training dynamics of bilingual language models. *3rd Multilingual Representation Learning Workshop* (workshop at EMNLP). Extended abstract.

Chang, T. A., & Rafferty, A. N. (2020). Encodings of source syntax: Similarities in NMT representations across target languages. *Proceedings of the 5th Workshop on Representation Learning for NLP* (workshop at ACL).

OTHER WORK / PREPRINTS

Chang, T. A., Tu, Z., & Bergen, B. K. (under review). Learning, forgetting, and stability: Characterizing learning curves during language model pre-training.

Chang, T. A., Arnett, C., Tu, Z., & Bergen, B. K. (under review). When is multilinguality a curse? Language modeling for 250 high- and low-resource languages.

Chang, T. A. (2020). Emergence of hierarchical syntax in neural machine translation. *Carleton Digital Commons*. Undergraduate thesis, Carleton College Cognitive Science. With distinction.

Chang, T. A. (2020). Topology of second order tensor fields. *Carleton Digital Commons*. Undergraduate thesis, Carleton College Mathematics and Statistics.

TEACHING

Teaching Assistant, UC San Diego 2021
  COGS188 AI Algorithms, Winter 2021
  COGS108 Data Science in Practice, Fall 2021

Course Staff, Carleton College 2019
  CS254 Computability and Complexity, Winter 2019
  CS111 Introduction to Computer Science, Spring 2019

AWARDS AND FELLOWSHIPS

Graduate Prize Fellowship, Halıcıoğlu Data Science Institute	2020-2024
Glushko Travel and Research Award, UCSD Cognitive Science	2020-2024
Roy O. Elveton Prize in Cognitive Science and Philosophy, Carleton College	2020
Google Spot Bonus, Google Geo Machine Perception	2019
Patricia V. Damon Scholarship, Carleton College	2019-2020
Phi Beta Kappa Third-Year Inductee, Carleton College	2019
Phi Beta Kappa First Year Prize, Carleton College	2018
Dean's List, Carleton College	2017, 2018, 2019
Stuebe Endowed Scholarship, Carleton College	2017-2020
National Merit Scholar, National Merit Scholarship Corporation	2016-2020

SERVICE

ACL Rolling Review (ARR) reviewer (Association for Computational Linguistics)	2023-2024
BMC Medical Education reviewer (BioMed Central)	2024
CogSci reviewer (Conference of the Cognitive Science Society)	2022, 2023, 2024
ACL reviewer (Annual Meeting of the Association for Computational Linguistics)	2023
FAccT reviewer (Conference on Fairness, Accountability, and Transparency)	2023
EMNLP reviewer (Conference on Empirical Methods in NLP)	2022, 2023
PNAS reviewer (Proceedings of the National Academy of Sciences)	2022
UCSD Graduate Application Mentorship Program, mentor	2022-2023
ICLR volunteer (International Conference on Learning Representations)	2022
ACL conference volunteer (Association for Computational Linguistics)	2020, 2021, 2022
EMNLP conference volunteer (Empirical Methods in NLP)	2020, 2022
Carleton Alumni Relations Office, student supervisor	2017-2020
Carleton Alumni Relations Office, student caller	2016-2017