Tyler A. Chang

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

I am a cognitive science PhD candidate at UC San Diego working on the analysis of large language models, particularly during pretraining. Most recently, I spent two years as a student researcher at Google DeepMind and Google Research, researching hallucination detection and scaling influence functions to LLM pretraining. Previously, I've interned at Amazon AI Labs, Alexa AI, Google Cloud Translate, and Google Earth. My research aims to enable more transparent, inclusive, and auditable language technologies.

EDUCATION

2020-Pres.	**University of California San Diego**, La Jolla, CA PhD Candidate in Cognitive Science Halıcıoğlu Data Science Institute, Graduate Fellow
2016-2020	**Carleton College**, Northfield, MN BA in Mathematics, summa cum laude BA in Cognitive Science, with distinction
2018	**Eötvös Loránd University**, Budapest, Hungary Budapest Semesters in Mathematics
2018	**Doshisha University**, Kyoto, Japan Carleton Linguistics and Culture Program

RESEARCH FOCUS

Language model analysis: pretraining dynamics, influence functions, multilinguality.
Cognitive linguistics: distributional semantics, language acquisition.

INDUSTRY

2022-2024	**Google DeepMind and Google Research**, PhD Student Researcher Pretraining, tuning, and analysis of large language models. Trained large language models from scratch, analysing petabytes of model activations and gradients.
2021, 2022	**Amazon Science**, Applied Scientist Intern AWS AI Labs: linguistic dataset drift and out-of-domain model generalization. Alexa AI: content change prediction on the web for intelligent web crawling.
2019, 2020	**Google**, Software Engineering Intern Cloud AI Translation: automatic sentence pair extraction from translated text. Geo Machine Perception (Maps, Earth): aerial and street level imagery alignment at scale.
2018	**Google**, Engineering Practicum Intern Chrome Web Store: built servers for tens of millions of monthly users.

PREPRINTS UNDER REVIEW

2024 **Chang, T. A.**, Rajagopal, D., Bolukbasi, T., Dixon, L., & Tenney, I. Scalable influence and fact tracing for large language model pretraining.

2024 **Chang, T. A.**, Arnett, C., Tu, Z., & Bergen, B. K. Goldfish: Monolingual language models for 350 languages.

PUBLICATIONS

Peer-reviewed.

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

2024 **Chang, T. A.**, Tu, Z., & Bergen, B. K. Characterizing learning curves during language model pre-training: Learning, forgetting, and stability. *Transactions of the Association for Computational Linguistics* (TACL).

2024 **Chang, T. A.**, Arnett, C., Tu, Z., & Bergen, B. K. When is multilinguality a curse? Language modeling for 250 high- and low-resource languages. *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP). Received an outstanding paper award.

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

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

2024 **Chang, T. A.***, Tomanek, K.*, Hoffmann, J., Thain, N., van Liemt, E., Meier-Hellstern, K., & Dixon, L. 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.

2024 Shah, C.♢, Chandak, Y.♢, Mane, A.♢, Bergen, B. K., & **Chang, T. A.** 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). ♢Undergraduate mentees.

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

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

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

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

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

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

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

2021 **Chang, T. A.**, Xu, Y., Xu, W., & Tu, Z. 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).

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

WORKSHOP PAPERS AND ABSTRACTS

Peer-reviewed.

2024 Arnett, C.*, **Chang, T. A.***, & Bergen, B. K. 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.

2024 Arnett, C.*, Rivière, P. D.*, **Chang, T. A.**, & Trott, S. Different tokenization schemes lead to comparable performance in Spanish number agreement. *Proceedings of the SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology* (workshop at NAACL). *Equal contribution.

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

2020 **Chang, T. A.**, & Rafferty, A. N. 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

2024 **Chang, T. A.**, Rajagopal, D., Bolukbasi, T., Dixon, L., & Tenney, I. Scaling training data attribution: The science of how training data influences LLM behavior. *People & AI Research, Google DeepMind, Medium Blog*.

2023 Thain, N., **Chang, T. A.**, Dixon, L., Croak, M., Meier-Hellstern, K., Tomanek, K., Hoffmann, J., & van Liemt, E. Detecting errors in chat bot outputs using language model neural networks. *United States patent application*. Work done at Google Research.

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

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

TEACHING

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

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

AWARDS AND FELLOWSHIPS

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

SERVICE

Reviewer Proceedings of the National Academy of Sciences (PNAS)
Nature Scientific Reports (Springer Nature)
Journal of Memory and Language (JML)
Artificial Intelligence Review (Springer)
EPJ Data Science (European Physical Journal)
BMC Medical Education (BioMed Central)

ACL Rolling Review (ARR, Association for Computational Linguistics)
ICLR 2025 (International Conference on Learning Representations)
CogSci 2022-2024 (Conference of the Cognitive Science Society)
ACL 2023-2024 (Meeting of the Association for Computational Linguistics)
FAccT 2023 (Conference on Fairness, Accountability, and Transparency)
EMNLP 2022-2024 (Conference on Empirical Methods in NLP)

2022-2023 UCSD Graduate Application Mentorship Program, mentor
2022 ICLR volunteer (International Conference on Learning Representations)
2020-2022 ACL conference volunteer (Association for Computational Linguistics)
2020, 2022 EMNLP conference volunteer (Empirical Methods in NLP)