Yan Cong

📱 *+1 (517) 940 1452*
✉ *cong4@purdue.edu*
🌐 *https://yancong222.github.io/*

Professional positions

08/2023- **Assistant Professor**, *School of Languages and Cultures, Purdue University*

08/2023- **Affiliate**, *School of Interdisciplinary Studies, Linguistics, Purdue University*

07/2023 **Postdoctoral fellow**, *Feinstein Institutes, Northwell Health*
Specializations: Natural Language Processing, Speech and Language Pathology, Data engineering

---

Education

07/2021 **Ph.D. in Linguistics**, *Michigan State University*
Dissertation: *Competition in Natural Language Meaning: The Case of Adjective Constructions in Mandarin Chinese and Beyond*

02/2015 **M.A. in Language Studies**, *Hong Kong Baptist University*

Journal Articles and Book Chapters

2023 Tang, Sunny; **Cong, Yan**; Mercep, Gwenyth; Bhatti, Mutahira; Serpe, Grace; Gromova, Valeria; Berretta, Sarah; John, Majnu; Liberman, Mark; Sinvani, Liron. *Journal of Psychiatry and Neuroscience*. Characterizing and Detecting Delirium with Clinical and Computational measures of Speech and Language Disturbance. `https://doi.org/10.1503/jpn.230026`

2022 **Cong, Yan**. *Asian Language and Linguistics* 2:2, pp.167-216. Judge-dependence in Quality Nouns: A Semantic Analysis of the Mandarin Chinese "*you* NP" Structure. `https://doi.org/10.1075/alal.21007.con`

2022 Tang, Sunny; **Cong, Yan**; Nikzad, Amir; Mehta, Aarush; Cho, Sunghye; Hänsel, Katrin; Beretta, Sarah; Dhar, Aamina; Kane, John; Malhotra, Anil. *Special Issue of Schizophrenia Research*. Clinical and Computational Speech Measures are Associated with Social Cognition in Schizophrenia Spectrum Disorders. *Schizophrenia research*. `https://doi.org/10.1016/j.schres.2022.06.012`

2022 Nikzad, Amir; **Cong, Yan**; Beretta, Sarah; Hänsel, Katrin; Cho, Sunghye; Pradhan, Sameer; Behbehani, Leily; DeSouza, Danielle; Liberman, Mark; Tang, Sunny. *Nature Schizophrenia*. Who Does What to Whom? Graph Representations of Action-Predication in Speech Relate to Psychopathological Dimensions of Psychosis. *Schizophrenia 8, 58*. `https://doi.org/10.1038/s41537-022-00263-7`

2022 **Cong, Yan**; Ngonyani, Deo. A syntactic analysis of the co-occurrence of stative and passive in Kiswahili. In Sibanda, Galen; Ngonyani, Deo; Choti, Jonathan; Biersteker, Ann (eds.), *Descriptive and theoretical approaches to African linguistics*, 165–181. Berlin: Language Science Press. DOI: 10.5281/zenodo.6393748. `https://langsci-press.org/catalog/book/306`

2022 Tang, Sunny; Hänsel, Katrin; **Cong, Yan**; Nikzad, Amir; Cho, Sunghye; Berretta, Sarah; Behbehani, Leily. Latent Factors of Speech and Language Disturbance and Relationships to Acoustic and Lexical Computational Features. *Special Issue of Schizophrenia Bulletin*. `https://doi.org/10.1093/schbul/sbac145`

Peer-reviewed Proceedings Papers

2023 **Cong, Yan**; Chersoni, Emmanuele; Hsu, Yu-Yin; Blache, Philippe. *Investigating the Effect of Discourse Connectives on Transformer Surprisal: Language Models Understand Connectives, Even So They Are Surprised*. In Proceedings of the 6th Conference on Blackbox NLP, pages 222-232, Association for Computational Linguistics. `https://aclanthology.org/2023.blackboxnlp-1.17/`

2023 **Cong, Yan**; Chersoni, Emmanuele; Hsu, Yu-Yin; Lenci, Alessandro. *Are Language Models Sensitive to Semantic Attraction? A Study on Surprisal*. In Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023), pages 141-148, Association for Computational Linguistics. `https://aclanthology.org/2023.starsem-1.13.pdf`

2022 **Cong, Yan**. *Psycholinguistic Diagnosis of Language Models' Commonsense Reasoning*. In Proceedings of the First Workshop on Commonsense Representation and Reasoning (CSRR 2022), pages 17–22, Dublin, Ireland. Association for Computational Linguistics. `https://aclanthology.org/2022.csrr-1.3/`

2022 **Cong, Yan**. *Pre-trained Language Models' Interpretation of Evaluativity Implicature: Evidence from Gradable Adjectives Usage in Context*. In Proceedings of the Second Workshop on Understanding Implicit and Underspecified Language, pages 1–7, Seattle, USA. Association for Computational Linguistics. `https://aclanthology.org/2022.unimplicit-1.1/`

2021 Pandia, Lalchand; **Cong, Yan**; Ettinger, Allyson. *Pragmatic competence of pre-trained language models through the lens of discourse connectives*. In Proceedings of the 25th Conference on Computational Natural Language Learning, pages 367–379, Online. Association for Computational Linguistics. `https://aclanthology.org/2021.conll-1.29/`

2019 **Cong, Yan**. *The Non-canonical Argument Realization of Chinese Alternations*. In Proceedings of the 30th North American Conference on Chinese Linguistics (NACCL-30). Volume 1. Pages 82-99. `https://www.researchgate.net/publication/341598112_The_Non-canonical_Argument_Realization_of_Chinese_Alternations`

2015 **Cong, Yan**; Wee, Lian-Hee; Huang, Chu-Ren. *The Invertible Construction in Chinese*. In Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation, pages 206–211, Shanghai, China. `https://aclanthology.org/Y15-1024/`

Invited talks

2023 **Cong, Yan**. *Large language models and their implications in text analysis*. Round table on *STEM in Asia*. Department of History, Purdue University.

2023 **Cong, Yan**. *Exploring L2 Writing Proficiency via LLMs Metrics: a Preliminary Study*. Invited talk for *LINGTalk*. Department of Linguistics, Purdue University.

2023 **Cong, Yan**. *Using large language models to measure fluency*. Advanced Methods at Purdue (AMAP) in the behavioral, health, and social sciences. Purdue University.

2023 **Cong, Yan**; Nikzad, Amir; Tang, Sunny X. *Multi-modal computational speech and language features for assessing psychiatric conditions*. The University of Maryland.

2023 **Cong, Yan**; Hänsel, Katrin; Nikzad, Amir; Tang, Sunny X. *Investigating novel NLP approaches for measuring speech disturbances in psychosis*. The Functional Neuroimaging and Bioinformatics Lab. McLean, Harvard Medical School Affiliate.

2022 **Cong, Yan** and Tang, Sunny X. *The 6th Annual Technology in Psychiatry Summit.* Leveraging state-of-the-art transformer language models to measure speech disturbance in schizophrenia. Harvard Medical School, Mclean Hospital.

2021 **Cong, Yan**. *Formal Linguistics and Philosophy (FLaP)*. What we don't say matters. The Hong Kong Baptist University, Department of English.

Conference Presentations and Peer-Reviewed Abstracts (Selected)

2024 Issa, Elsayed; **Cong, Yan**; Scala, Carmela; Merschel, Lisa; Frattarola, Angela. Organizers: William, Jennifer; Fellin, Luciana. AI, NLP, and Computational Linguistics: Implications for Language Programs. Accepted to *Modern Language Association (MLA) 2024 Convention* (January 4-7, 2024; Philadelphia, PA).

2023 Nikzad, Amir H.; **Cong, Yan**; Berretta, Sarah; Behbehani, Leily; Pradhan, Sameer; Tang, Sunny X.. Disorganization Disrupts the Core: Localizing Clinical Dimensions of Speech and Language Disturbance Within Clause Structure Using Multi-Layered Semantic Graph Analysis. Accepted to *the 62nd Annual Meeting of the American College of Neuropsychopharmacology* (December 3-6, 2023; Tampa, FL).

2023 Tang, Sunny X.; Nikzad, Amir H.; Spilka, MJ; **Cong, Yan**; Behbehani, Leily; Berretta, Sarah; Kane, JM; Malhotra, AK. Linguistic risk score for schizophrenia: applying a polygenic risk approach to synthesize multivariate features from computational speech and language analysis. Accepted to *the 62nd Annual Meeting of the American College of Neuropsychopharmacology* (December 3-6, 2023; Tampa, FL).

2022 **Cong, Yan**; Wolff, Phillip. *Annual Meeting of The Linguistic Society of America (LSA)*. Inferring markedness from semantic weight: An approach using the T5 language model. Washington, D.C.

2022 Nikzad, Amir; **Cong, Yan**; Hänsel, Katrin; Cho, Sunghye; Berretta, Sarah; Behbehani, Leily; Tang, Sunny. *Annual Meeting of The Society of Biological Psychiatry (SOBP)*. Modeling cross-diagnostic speech disturbance with semantic and structural speech graphs. Virtual presentation.

2022 Hänsel, Katrin; Nikzad, Amir; **Cong, Yan**; Mehta, Aarush; Cho, Sunghye; Berretta, Sarah; Behbehani, Leily; Tang, Sunny. *Annual Meeting of The Society of Biological Psychiatry (SOBP)*. Predicting general psychopathology with multi-modal speech and language features. Virtual presentation.

2021 **Cong, Yan**; Buccola, Brian. *The 45th annual Penn Linguistics Conference (PLC)*. Alternatives and symmetry in non-scalar implicature: The case of Mandarin adjective constructions. Philadelphia, Pennsylvania.

2021 Hänsel, Katrin; Cho, Sunghye; Berretta, Sarah; Mehta, Aarush; Nikzad, Amir; Guo, Jessica; **Cong, Yan**; Pradhan, Sameer; Fiumara, James; Liberman, Mark; Tang, Sunny. *The 60th Annual meeting of American College of Neuropsychopharmacology (ACNP)*. Associative dysregulation and impaired expressivity in schizophrenia: Latent structures of clinical and automated speech analysis features. Virtual presentation.

2019 Smith, Kaylin; Ma, Ye; **Cong, Yan**; Beretta, Alan. *The Psycholinguistics in Iceland— Parsing and Prediction (PiPP)*. Interactions between syntax and semantics: The semantic P600, University of Iceland. Reykjavík, Iceland.

2019 Smith, Kaylin; Ma, Ye; **Cong, Yan**; Beretta, Alan. *The 32nd CUNY Conference on Human Sentence Processing*. Semantic attraction in sentence processing. University of Colorado Boulder. Boulder, Colorado.

2019 **Cong, Yan**. *The Linguistic Society of America Linguistics Institute (LSA)*. Information structure and the depth of language processing: a pilot study. University of California at Davis. California.

2019 **Cong, Yan**. *International Workshop on Degrees and Grammar: An East Asian Perspective*. A subjective analysis of the Mandarin '*you*' structure. Nanjing University, China.

2018 **Cong, Yan**. *The Annual Conference of the Chinese Language Teachers Association (CLTA)*. The non-canonical argument realization of the Chinese alternations. Washington, D.C.

2018 **Cong, Yan**. *The 12th Arizona Linguistics Circle (ALC-12)*. Rethinking the implicit argument in Kiswahili. University of Arizona. Tucson, Arizona.

2018 **Cong, Yan**. *The 2nd Crete Linguistics Summer School (CreteLing-2)*. A preliminary semantic analysis of predicates of taste in Chinese. University of Crete. Greece.

2017 Parrish, Alicia; Kelley, Patrick; Smith, Kaylin; **Cong, Yan**; Beretta, Alan. *Annual Meeting of The Society for the Neurobiology of Language (SNL)*. Prosodic lengthening and boundary prediction in nominal compounds: An ERP study. Baltimore, Maryland.

2017 **Cong, Yan**. *The Linguistic Society of America Linguistics Institute (LSA)*. T1 and T0 Sandhi in Tongliao Mandarin: an OT approach. University of Kentucky. Lexington, Kentucky.

2016 **Cong, Yan**. *The 28th North American Conference on Chinese Linguistics (NACCL-28)*. Markedness theory and the acquisition of the Mandarin potential complement construction. Brigham Young University. Salt Lake City, Utah.

Fellowships, Funds, and Awards

2023 **Internal fund**, ASPIRE research enhancement grant.
Travel for *Modern Language Association (MLA) Annual Convention*. Purdue CLA $1,550

2023 **Internal fund**, Start-up funds.
Purdue CLA $40,000

2023 **Educational Advancement Award**, *Advancing Women in Science and Medicine (AWSM)*
Computational approach to speech biomarkers in bilingual immigrants with schizophrenia.
Northwell health $5,000

2022 **Educational Advancement Award**, *Advancing Women in Science and Medicine (AWSM)*
Computational linguistics can predict psychiatric diagnoses and outcome.
Northwell health $5,000

2020 **Dissertation Completion Fellowship**
MSU Graduate School $10,000

06/2018 **Student Fellowship**, *North American Summer School on Logic, Language, and Information (NASSLLI)*, at Pittsburgh, Pennsylvania
Carnegie Mellon University $500

2017-2020 **Summer Support Fellowship**
MSU Department of Linguistics and Languages $4,500

2015 **Best Paper Award**, *The 29th Pacific Asia Conference on Language, Information and Computation*, The invertible construction in Chinese. Shanghai Jiaotong University, China

Teaching

Undergrad
- SCLA102: Cornerstone - Critical thinking and communication, Purdue (2023-present)
- LIN401: Introduction to Linguistics, MSU (2017-2021)
- CHS101, CHS102, CHS411: Chinese language, MSU (2016-2017)

Grad
- LING598-LC596: Statistical Modeling in Linguistics, Purdue (2023)

Advising

01/2024 – present
Ph.D committee member (in progress)
- Yue Li. *A Project on English Object Relative Clauses Extraction and Categorization Using Python*

Service to the profession

Reviewer	Journal of Medical Internet Research	2024-present
Reviewer	Computer Speech & Language	2023-present
Reviewer	Scientific Reports	2023-present
Reviewer	Health Informatics Journal	2023-present
Reviewer	The Association for Computational Linguistics	2022-present
Reviewer	Association of Contemporary African Linguistics	2020-2021

Service to Purdue University

Member	Graduate admissions subcommittee.	for Fall 2024
Member	CLA Ross-Lynn Postdoctoral Fellow review subcommittee.	for Fall 2024
Member	Purdue Computational Linguistics Lab.	2023-present

Associations

Member	Modern Language Association	2023-present
Member	New York Academy of Sciences	01/2022-07/2023
Member	Society for the Neurobiology of Language	2018-present
Member	The Chinese Language Teachers Association – USA (CLTA-USA)	2018-2019

January 15, 2024