

Xixian Liao

CONTACT INFORMATION

🌐 xixianliao.github.io
@ xixianliao@gmail.com

WORK EXPERIENCE

Research Engineer, Language Technologies Laboratory at Barcelona Supercomputing Center 04/2024–
Chinese-Spanish Interpreter & Translator, China’s Ministry of Commerce 2017–2018

EDUCATION

Ph.D. in Computational Linguistics (*cum laude*), Universitat Pompeu Fabra, Spain 2019–2024
Visiting graduate student, University of Edinburgh, UK 2022
M.A. in Theoretical and Applied Linguistics (GPA: 8.68/10), Universitat Pompeu Fabra, Spain 2018–2019
B.A. in Hispanic Philology (GPA: 9.58/10, with honors), Beijing Jiaotong University, China 2014–2018

ARTICLES IN SCIENTIFIC JOURNALS

Liao, X., Brochhagen, T., Boleda, G., and Mayol, L. In press. The role of referent predictability in pronoun production: Insights from a Bayesian meta-analysis (pre-print). *Language*. [data and code]
Liao, X., Boleda, G., Rohde, H., and Mayol, L. (2024). Comparing models of pronoun production and interpretation via observational and experimental evidence. *Glossa: a journal of general linguistics* 9(1). doi: <https://doi.org/10.16995/glossa.10884> [data and code]

ARTICLES IN PROCEEDINGS

Liu, M., Dong, Y., **Liao, X.**, and Webber, B. (2025). Multi-token Mask-filling and Implicit Discourse Relations. *Findings of EMNLP*.
Liao, X., Escolano, C., Mash, A., De Luca Fornaciari, F., García Gilabert, J., Claramunt Argote, M., Bohman, E., and Melero, M. (2025). Culture-aware machine translation: The case study of low-resource language pair Catalan-Chinese. *Proceedings of Machine Translation Summit XX*, 150–161.
García Gilabert, J., Escolano, C., Sant, A., De Luca Fornaciari, F., Mash, A., **Liao, X.**, and Melero, M. (2025). Investigating the translation capabilities of large language models trained on parallel data only. *Proceedings of Machine Translation Summit XX*, 24–53.
García Gilabert, J., Escolano, C., Mash, A., **Liao, X.**, and Melero, M. (2025). MT-LENS: An all-in-one toolkit for better machine translation evaluation. *Proceedings of NAACL 2025: Human Language Technologies (System Demonstrations)*, 51–60.
Sant, A., Bardanca, D., Pichel Campos, J. R., De Luca Fornaciari, F., Escolano, C., García Gilabert, J., Gamallo, P., Mash, A., **Liao, X.**, and Melero, M. (2024). Training and fine-tuning NMT models for low-resource languages using Apertium-based synthetic corpora. *Proceedings of the Ninth Conference on Machine Translation*, 925–933.
Liang, J. and **Liao, X.** (2024). How does grammatical category influence conceptual categorization: The case of Chinese classifiers. *Proceedings of CogSci*, 46.
Dong, Y., **Liao, X.**, and Webber, B. (2024). Syntactic preposing and discourse relations. *Proceedings of EACL*, 2790–2802.
He, Y., **Liao, X.**, Liang, J., and Boleda, G. (2023). The impact of familiarity on naming variation: A study on object naming in Mandarin Chinese. *Proceedings of CoNLL*, 456–475.

- Aina*, L., Liao*, X., Boleda, G., and Westera, M. (2021). Does referent predictability affect the choice of referential form? A computational approach using masked coreference resolution. *Proceedings of CoNLL*, 454–469.
- Liao, X. Coherence-driven predictability and referential form: evidence from English corpus data. *Proceedings of Sinn und Bedeutung* 26, 544–556.

FELLOWSHIPS, SCHOLARSHIPS, GRANTS

PhD fellowship FI SDUR grant (competitive Ph.D. fellowship), AGAUR, Spain	2020–2023
CSC scholarship (competitive Ph.D. fellowship), Chinese Scholarship Council	2019–2020

SUPERVISION

MASTER’S THESES

Co-supervisor, Chelsea Chuang, 2025, MSc in Speech and Language Processing, The University of Edinburgh.

Co-supervisor, Tianyi Li, 2025, MSc in Data Science, The University of Edinburgh.

Co-supervisor, Nick Zhu, 2025, MSc in Cognitive Science, The University of Edinburgh.

Co-supervisor, Meinan Liu, 2024, MSc in Cognitive Science, The University of Edinburgh.

Co-supervisor, Zizhe Wang, 2024, MSc in Artificial Intelligence, The University of Edinburgh.

Co-supervisor, Yunfang Dong, 2023, MSc in Speech and Language Processing, The University of Edinburgh.

Co-supervisor, Yunke He, 2023, MSc in Theoretical and Applied Linguistics, Universitat Pompeu Fabra.

TEACHING & ACADEMIC SERVICE

TEACHING

Universitat Pompeu Fabra

<i>Computational Semantics</i> , Master’s Theoretical and Applied Linguistics (Teaching Assistant)	2023
<i>Information and Communications Technology</i> , B.A. Applied Languages (Teaching Assistant)	2020–2023
<i>Semantics and Pragmatics</i> , Master’s Theoretical and Applied Linguistics (Teaching Assistant)	2021

REVIEWING

Journals: *Journal of Memory and Language*; *Language, Cognition and Neuroscience*

Conferences and workshops: ACL ARR; *SEM 2025; WMT 2025

OTHER SERVICE

Area Chair, CoNLL 2025	2025
Co-chair (track: Language and Computation), Student Session of the 34th European Summer	2022
School in Logic, Language, and Information, National University of Ireland Galway	
Co-organizer, REtreat for Students in Computational Linguistics (REST-CL), Spain	2022–23
Co-organizer, GLiF research group seminar, Universitat Pompeu Fabra	2020–2021

PRESENTATIONS & TALKS

CONFERENCE AND WORKSHOP PRESENTATIONS

- Culture-aware machine translation: the case study of low-resource language pair Catalan-Chinese. *Machine Translation Summit*. Geneva, Switzerland, 23–27 June, 2025.
- The Impact of Familiarity on Naming Variation: A Study on Object Naming in Mandarin Chinese. *The SIGNLL Conference on Computational Natural Language Learning (CoNLL)*. Singapore, Dec 6–7, 2023.
- Bayesian pronoun interpretation in English corpus passage completions. *36th Annual Conference on Human Sentence Processing*. University of Pittsburgh. March 9–11, 2023.

Does referent predictability affect the choice of referential form? A computational approach using masked coreference resolution. *The SIGNLL Conference on Computational Natural Language Learning (CoNLL)*. Virtually, Nov 10–11, 2021; *Deep Learning Barcelona Symposium*. Universitat Politècnica de Catalunya, Dec 22–23, 2021.

Coherence-driven predictability and referential form: evidence from English corpus data. *Sinn und Bedeutung 26*. University of Cologne, Sept 8–10, 2021; *Student session in European Summer School in Logic, Language and Information (ESSLLI)*. Virtually, Aug 4–13, 2021.

TALKS

Bayesian Meta-Analysis Methods. *Life Sciences Department Seminar*. Barcelona Supercomputing Center, April 10, 2025.

SalamandraTA Machine Translation. *Bioinfo4Women*. Life Sciences Open Day at Barcelona Supercomputing Center, February 11, 2025.

Predictability and pronoun production: an online passage continuation study and a Bayesian meta-analysis (two studies in progress). *PSYCHOLINGUISTICS COFFEE*. University of Edinburgh, June 15, 2022.

OTHER DISCUSSION

Linguistics and Language Models: What Can They Learn from Each Other? *Dagstuhl Seminar 25301*. Schloss Dagstuhl, Germany, July 20–25, 2025.

LANGUAGES & SKILLS

LANGUAGES

Mandarin Chinese (native), English (IELTS C1), Spanish (DELE C1), Catalan (A2)

SKILLS

Python, R, HTML, Java