

Xiaochuan Li

Affiliation:	M.Sc. student Human and Artificial Intelligence University of Technology Nuremberg	Homepage:	xchuan-li.github.io
E-mail:	xiaochuan.li@utn.de	GitHub:	github.com/xchuan-li

What is there in language itself that we have not yet discovered? Linguistics and cognitive science offer precise questions about meaning and decades of measured human behaviour. In language models, those questions become controlled experiments: training shows what can be learned from language alone, behavioural tests show what a model does, and interpretability shows where in the model it happens.

Education

Oct. 2025 – exp. 2027	University of Technology Nuremberg , Germany <i>M.Sc. Human and Artificial Intelligence</i> Designing and analysing behavioural experiments; cognitive psychology; training neural networks; interpretability methods for language models.
Oct. 2024 – Jul. 2025	University of Marburg , Germany <i>German for graduate study</i>
Sep. 2020 – Jun. 2024	Shenzhen University , China <i>B.A. Philosophy</i> Formal logic and proof, including modal logic and Gödel’s incompleteness theorems; philosophy of language; psychology of reading.

Research

Apr. 2026 – Mar. 2027	Moral framing and dilemma choices in language models <i>Supervised at the German Research Center for Artificial Intelligence (DFKI)</i> Investigated how moral wording affects dilemma choices. Designed and ran controlled experiments and activation interventions in 7B and 14B models for the utilitarian-framing branch, with checks of answer labels, option positions, and task understanding; manuscript in preparation.
Sep. 2026 –	Feature dependency and property answers for modified nouns in language models <i>Independent project</i> Contrasted properties that depend on an added feature with properties that only co-occur with it (13 feature groups, 477 items, three instruction-tuned models), using predictors fixed outside the models: a dependency table, human feature norms and word-association norms; exploratory activation patching. Manuscript in preparation.
Oct. 2026 –	Vagueness and learning a scale from text <i>Independent project</i> A proposition on when speakers’ disagreement about word boundaries, a small cost to a listener, gives a text-only learner evidence about the scale, tested with language models trained from scratch on naming text (artificial language; World Color Survey). Preprint. doi:10.5281/zenodo.23677586

Sep. 2026 – **Modal language in preschool children**
M.Sc. thesis, design stage · part of a PhD project at the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig
 Asks how preschool children come to distinguish possibility from necessity in language, and how this relates to their nonlinguistic reasoning about possibilities; a behavioural study; data collection has not started. It lays the groundwork for later work on modal language in language models.

Projects

Sep. 2026 Typicality in referent choice: what the description leaves open
Independent theoretical study. Distinguishes three conditions for typicality in referent choice and the limits of what a forced-choice task can establish. Li, X. (2026). Working paper. lingbuzz/010343 · [doi:10.5281/zenodo.22855036](https://doi.org/10.5281/zenodo.22855036)

May – Aug. 2026 Cross-lingual ProtoBias
Single-author course project. Evaluated two VLM families on fixed image pairs across seven prompt languages (about 12,600 judgments); single-author report with code. [report](#) · [code](#)

May – Jul. 2026 Street-view country classification
Team course project, Deep Learning. My part: CNN training, ablations, and performance analysis.

Nov. 2025 – Feb. 2026 Fact-checking after ChatGPT exposure
Course study, Foundations in Psychology and Empirical Study Design. Between-subjects experiment on optional checking after ChatGPT or Wikipedia answers (49 participants, 47 analysed); co-designed the study, ran the data analysis and wrote the report.

Dec. 2023 – May 2024 Leibniz's logical system and *characteristica universalis*
Undergraduate thesis · Shenzhen University. Analysed Leibniz's metaphysics, logic and proposal for a universal symbolic language.

Employment

Apr. 2026 – present **University of Technology Nuremberg, Germany**
IT Service Desk (part-time)
 Delivered student IT onboarding and individual support; handled system installation, service reporting, and inventory.

Sep. – Dec. 2021 **Shenzhen Learning Disorders Association, China**
Internship · Learning assessment at Yucai No. 2 Primary School
 Prepared character-reading materials, administered timed assessments, and analysed response data to support the assessment of learning difficulties; contributed to teaching items and individual learning plans.

Skills and languages

Language models	Training small models from scratch; controlled behavioural evaluation, minimal pairs; activation patching, low-rank analysis, probing, steering
Human data	Experimental design, PsychoPy / Pavlovia, JASP; testing children
Programming	Python (PyTorch, Hugging Face Transformers, pandas, scikit-learn), Git, LaTeX, Slurm on an HPC cluster
Languages	Mandarin Chinese (native), German (C1), English (C1), Japanese (learning)