

徐炆个人简历

2026

单位：南方科技大学，计算机科学与工程系
职称：副教授

邮箱: sustech@sustech.edu.cn
主页: [clcs-sustech.github.io](https://github.com/clcs-sustech)

研究方向

徐炆，南方科技大学副教授，博士生导师。美国宾夕法尼亚州立大学博士，清华大学硕士和学士。广东省珠江人才计划青年拔尖人才，深圳市孔雀计划人才（C类）。

徐炆博士的研究方向集中在计算语言学、自然语言处理、心理语言学和认知科学的交叉领域。致力于探索人类自然语言现象与内在认知、心理活动，以及心智意识现象的关联。目前主持南方科技大学计算语言学与意识科学（CLCS）实验室，研究团队主要关注：

- 人类自然语言与大型语言模型（LLMs）生成的区别。
- 语言现象对人脑认知活动的反映。
- 心理学与认知科学对人工智能技术的增强。

工作经历

副教授	2023 年 6 月 – 至今
南方科技大学（SUSTech），计算机科学与工程系	
助理教授	2018 年 8 月 – 2023 年 5 月
美国圣地亚哥州立大学（San Diego State University），计算机科学系	

教育背景

博士	2013 年 8 月 – 2018 年 5 月
美国宾夕法尼亚州立大学（Penn State），信息科学与技术学院	
硕士	2010 年 9 月 – 2013 年 7 月
清华大学，心理学系	
工学学士	2006 年 9 月 – 2010 年 7 月
清华大学，电子工程系	

科研项目

《基于信息熵理论的大语言模型幻觉问题研究》
项目负责人，深圳市自然科学基金面上项目，2024-2027

《基于熵的大语言模型生成评估》（已结题）
项目负责人，南方科技大学-华为可信智能系统创新实验室校企横向项目，2025

PI, NSF-CRII-HCC (Award ID: 2105192), \$166K, 2021 – 2024
Modeling Computer-Mediated Task-Oriented Dialogues with Multi-Modality Information Theoretic Approaches

Co-PI, NIH (subaward to SDSU HealthLink), \$50K, 2021 – 2022

An NLP Approach to Analyze the Past and Present, and Inform the Future of Health Disparities Research

发表论文

1. Yu Wang, Leyi Lao, Langchu Huang, Gabriel Skantze, **Yang Xu**, and Hendrik Buschmeier. Investigating the representation of backchannels and fillers in fine-tuned language models. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026)*, 2026. (accepted; CCF-A)
2. Zhichen Liu, Yongyuan Li, and **Yang Xu**. Think in sentences: Explicit sentence boundaries enhance language model’s capabilities. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026)*, 2026. (accepted; CCF-A)
3. Yulin Ou, Yu Wang, **Yang Xu**, and Hendrik Buschmeier. Identifying the periodicity of information in natural language. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026)*, 2026. (accepted; CCF-A)
4. Hao An, Yibin Lou, Jiayi Guo, and **Yang Xu**. Purging the gray zone: Latent-geometric denoising for precise knowledge boundary awareness. In *Findings of the Association for Computational Linguistics: ACL 2026*, 2026. (accepted; CCF-A)
5. Zhichen Liu, Kaitong Qin, Linhan He, and **Yang Xu**. Wavedetect: Robust framework for machine-generated text detection via wavelet transform. In *Findings of the Association for Computational Linguistics: ACL 2026*, 2026. (accepted; CCF-A)
6. Jianjie Zheng, Zhichen Liu, Zhanyu Shen, Jingxiang Qu, Guanhua Chen, Yile Wang, **Yang Xu**, Yang Liu, and Sijie Cheng. Evaluating memory capability in continuous lifelog scenario. In *Findings of the Association for Computational Linguistics: ACL 2026*, 2026. (accepted; CCF-A)
7. Xiaoshan He, Xiaoqun Liu, Haodong He, Yu Wang, and **Yang Xu**. Eye movement features can predict human preferences on machine-generated texts. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop)*, 2026. (accepted)
8. Hao An and **Yang Xu**. Teaching llms to abstain via fine-grained semantic confidence reward. *arXiv*, 2026. (under review)
9. Zhichen Liu, Yongyuan Li, **Yang Xu**, Yu Wang, Yingfang Yuan, and Zuhao Yang. Evaluating text generation quality using spectral distances of surprisal. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, 2025
10. Zhongyu Yang, Junhao Song, Zhang Luo, Zuhao Yang, **Yang Xu**, Jingfen Lan, Yonghan Zhang, Wei Pang, Siyang Song, and Yingfang Yuan. Rechar: Revitalising characters with structure preserved and user-specified aesthetic enhancements. In *Proceedings of the SIGGRAPH Asia 2025 Technical Communications*, SA Technical Communications ’25. Association for Computing Machinery, 2025
11. Lam Nguyen and **Yang Xu**. Reasoning for translation: Comparative analysis of chain-of-thought and tree-of-thought prompting for LLM translation. In *Proceedings of the 63rd Annual Meeting of the Association*

for *Computational Linguistics (Volume 4: Student Research Workshop)*, pages 259–275, Vienna, Austria, July 2025. Association for Computational Linguistics

12. **Yang Xu**, Yu Wang, Hao An, Zhichen Liu, and Yongyuan Li. Detecting subtle differences between human and model languages using spectrum of relative likelihood. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 10108–10121, Miami, Florida, USA, November 2024. Association for Computational Linguistics. (CCF-B)
13. Yu Wang, **Yang Xu**, Gabriel Skantze, and Hendrik Buschmeier. How much does nonverbal communication conform to entropy rate constancy?: A case study on listener gaze in interaction. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 3533–3545. Association for Computational Linguistics, 2024. (CCF-A)
14. Zuhao Yang, Yingfang Yuan, **Yang Xu**, Shuo Zhan, Huajun Bai, and Kefan Chen. Face: Evaluating natural language generation with fourier analysis of cross-entropy. *Advances in Neural Information Processing Systems*, 36, 2024. (CCF-A, co-first author)
15. **Yang Xu** and Yang Cheng. Spontaneous gestures encoded by hand positions improve language models: An information-theoretic motivated study. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 9409–9424, 2023. (CCF-A)
16. **Yang Xu**, Yang Cheng, and Riya Bhatia. Gestures are used rationally: Information theoretic evidence from neural sequential models. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 134–140, Gyeongju, Republic of Korea, 2022. International Committee on Computational Linguistics. (CCF-B)
17. Bruce I. Gaynes, Mark B. Shapiro, Jie Xu, Abel Saju Augustine, **Yang Xu**, and et.al. Hierarchical data visualization of experimental erythrocyte aggregation employing cross correlation and optical flow applications. *Microvascular Research*, 143:104386, 2022
18. **Yang Xu**. Global divergence and local convergence of utterance semantic representations in dialogue. In *Proceedings of the Society for Computation in Linguistics*, volume 4, pages 116–124. 2021
19. **Yang Xu** and Zheng-sheng Zhang. Historical changes in semantic weights of sub-word units. In Adam Jatowt, Nina Tahmasebi, Yang Xu, and Lars Borin, editors, *Computational Approaches to Semantic Change*, pages 169–187. Language Science Press, 2021
20. M Alex Kelly, **Yang Xu**, Jesús Calvillo, and David Reitter. Which sentence embeddings and which layers encode syntactic structure? In *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*, pages 2375–2381, 2020. (acceptance rate: 18.1%; CCF-B)
21. **Yang Xu**, Jiasheng Zhang, and David Reitter. Treat the word as a whole or look inside? subword embeddings model language change and typology. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, pages 136–145, Florence, Italy, 2019. Association for Computational Linguistics
22. **Yang Xu**, Jeremy Cole, and David Reitter. Linguistic alignment is affected more by lexical surprisal rather than social power. *Proceedings of the Society for Computation in Linguistics*, 2(1):349–352, 2019
23. **Yang Xu**, Jeremy Cole, and David Reitter. Not that much power: Linguistic alignment is influenced

- more by low-level linguistic features rather than social power. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 601–610, 2018. (acceptance rate: 24.5%; CCF-A)
24. **Yang Xu** and David Reitter. Information density converges in dialogue: Towards an information-theoretic model. *Cognition*, 170:147–163, 2018. (Impact factor: 3.65; CCF-B)
 25. **Yang Xu** and David Reitter. Spectral analysis of information density in dialogue predicts collaborative task performance. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 623–633, Vancouver, Canada, 2017. (acceptance rate: 25%; CCF-A)
 26. **Yang Xu** and David Reitter. Entropy converges between dialogue participants: Explanations from an information-theoretic perspective. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 537–546, Berlin, Germany, August 2016. Association for Computational Linguistics. (acceptance rate: 28.0%; CCF-A)
 27. **Yang Xu** and David Reitter. Convergence of syntactic complexity in conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 443–448, Berlin, Germany, August 2016. Association for Computational Linguistics. (acceptance rate: 28.0%; CCF-A)
 28. **Yang Xu** and David Reitter. An evaluation and comparison of linguistic alignment measures. In *Proceedings of Cognitive Modeling and Computational Linguistics (CMCL)*, pages 58–67, Denver, CO, 2015. Association for Computational Linguistics
 29. Alexander G. Ororbia II, **Yang Xu**, Vito D’Orazio, and David Reitter. Error-correction and aggregation in crowd-sourcing of geopolitical incident information. In N. Agarwal et al., editor, *Social Computing, Behavioral Modeling and Prediction*, volume 9021 of *Lecture Notes in Computer Science*, pages 381–387. Springer, 2015
 30. David Reitter, **Yang Xu**, Patrick Craven, Anikó Sándor, R. Chris Garrett, E. Vince Cross, and Jerry L. Franke. Cognitive models predicting surprise in robot operators. In *Proc. International Conference on Cognitive Modeling*, pages 190–191, Groningen, Netherlands, 2015
 31. **Yang Xu** and Hong Li. Influence of visibility range and degree of urgency on efficiency of evacuation: Mediating effect of herding behavior. *Studies of Psychology and Behavior*, 13(3):311–319, 2015. (Published in Chinese)
 32. Hong Li, **Yang Xu**, Shi Chen, and Anqi Gao. Effects of intuition and analysis on the adoption of sources of information in escape decision-making. *Studies of Psychology and Behavior*, 10(6):452–458, 2012. (Published in Chinese)

会议报告

1. Global Divergence and Local Convergence of Dialogue, 线上, *SCiL*, 2021
2. Subword Embeddings Reveal Language Change, 海报, *CogSci*, 加拿大蒙特利尔, 2019
3. Treat the Word As a Whole or Look Inside, 海报, *ACL*, 意大利佛罗伦萨, 2019

特邀报告

加拿大滑铁卢大学	2022 年 8 月
邀请人: Moojan Ghafurian 博士	
题目: Speaking Rationally by Gestures: Evidence from Multi-Modal Language Models	
北京语言大学	2021 年 8 月
题目: The Information-Theoretic Insights of Multimodal Language Models	
西湖大学	2019 年 11 月
邀请方: 张岳博士课题组	题目: Dialogue from an Information-Theoretic Perspective
上海交通大学	2019 年 11 月
邀请方: 工学院	题目: Efficient Communication in Dialogue
清华大学	2019 年 11 月
邀请方: 心理学系	题目: An Information Theoretic Psycholinguistic Model of Dialogue
San Diego State University	2019 年 10 月
邀请方: SDSU 人工智能研讨会	题目: Recent Advances in Natural Language Processing
San Diego State University	2018 年 2 月
邀请方: 计算机系	题目: Computational Understanding of Dialogue and Language

荣誉与奖励

研究卓越奖	宾夕法尼亚州立大学信息科学与技术学院, 2018
优秀毕业生奖	清华大学, 2013
优秀硕士学位论文奖	清华大学, 2013

学术服务

ACL 2025–2026 领域主席 (Area Chair)
CogSci 2018–2020 审稿人
Lingua 审稿人, *Dialogue and Discourse* 审稿人, 2022

教学

@SUSTech	
CS310 Natural Language Processing, 自然语言处理 (本科生)	2024–2026
CS5026 Introduction and Frontiers of Cognitive Sciences, 认知科学导论与前沿 (研究生)	2025
CS208 Algorithms, 算法设计与实现 (本科生)	2023
@SDSU	
CS460 Algorithms	2018 – 2022
CS549 Machine Learning	2018 – 2022