

References

- Ahmad, A., Park, D., Dannenberg, A. L., Kiel, S., & Kearns, M. (2025). Distinguishing repetition disfluency from reduplication in spontaneous speech. *Proceedings of COLING 2025*. ACL Anthology.
- Alhammadi, W., Rabab'ah, G., & Alghazo, S. (2024). Gender differences in language use in Talks at Google. *KEMANUSIAAN: The Asian Journal of Humanities*, 31(1), 149–176. <https://doi.org/10.21315/kajh2024.31.1.8>
- Ariai, P., Mackenzie, J., & Demartini, G. (2024). AI and law: A survey of applications, challenges, and opportunities. *arXiv preprint arXiv:2403.12345*.
- Astuti, A. R. (2024). Speech disfluency and gestures production in undergraduate students. *JETLi: Journal of English Teaching and Linguistics*, 5(2), 72–83.
- Bajaj, A. K., Marone, M., & Mihalcea, R. (2023). Deception detection in conversations using proximity of linguistic markers. *Computers in Human Behavior: Artificial Humans*, 1, 100022.
- Barale, A., Rovatsos, M., & Bhuta, N. (2025). Fairness in legal NLP: Bias and transparency in automated adjudication. *AI and Law*.
- Belém, C. G., Kelly, M., Steyvers, M., Singh, S., & Smyth, P. (2024). Perceptions of linguistic uncertainty by language models and humans. *Proceedings of EMNLP 2024*. Association for Computational Linguistics.
- Benus, S., Enos, F., Hirschberg, J., & Shriberg, E. (2006). Pauses in deceptive speech. *Proceedings of Odyssey 2006: The Speaker and Language Recognition Workshop* (pp. 159–164). International Speech Communication Association.
- Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural correlates of disfluency in speech comprehension. *NeuroImage*, 109, 358–366. <https://doi.org/10.1016/j.neuroimage.2014.12.087>
- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, 44(2), 123–147.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Cameron, D. (1998). *The feminist critique of language: A reader*. Routledge.
- Chaudhry, A., Thiagarajan, S., & Gorur, D. (2024). Finetuning language models to emit linguistic expressions of uncertainty. *arXiv preprint arXiv:2409.12180*.

References

- Chen, J., Caesar, H., Tian, S., Li, B. Z., & Ortiz, L. E. (2020). Acoustic-prosodic and lexical cues to deception and trust in spoken dialogue. *Transactions of the Association for Computational Linguistics*, 8, 199–214.
- Chen, M., Pu, J., Shao, R., & Zhao, Y. (2023). Evaluating factual consistency in abstractive summarization. *Proceedings of ACL 2023*. Association for Computational Linguistics.
- Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., ... Le, Q. V. (2022). Scaling instruction-finetuned language models. *arXiv preprint arXiv:2210.11416*.
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111.
- Coates, J. (1996). *Women talk: Conversation between women friends*. Blackwell.
- Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, 2(4), 589–602.
- de Leeuw, E. (2014). Fluency in native and nonnative speech. *Language and Linguistics Compass*, 8(11), 581–589. <https://doi.org/10.1111/lnc3.12110>
- de Lima, L., Rezende, L., Consoli, R., & Kayser, M. (2024). Disfluency detection and removal in speech transcriptions via large language models. *Preprint*.
- Duncan, [Initials]. (2023). [Full reference needed — placeholder].
- Ehrlich, S. (2001). *Representing rape: Language and sexual consent*. Routledge.
- Escalona, M. (2025). Hedges and boosters in college essays: A study of linguistic modulation in student writing. *Journal of Language and Linguistic Studies*, 21(1), 1–21.
- Fang, Z., Liu, S., & Sun, M. (2025). Multi-model legal summarization. *Proceedings of ICAIL 2025*.
- Farzindar, A., & Lapalme, G. (2004). Legal text summarization: The case of Canadian immigration decisions. *Proceedings of the Workshop on Text Summarization*.
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2004). Good-enough representations in language comprehension. *Trends in Cognitive Sciences*, 8(1), 11–15.
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31(7), 931–952.
- Gal, Y. (2016). *Uncertainty in deep learning* (Doctoral dissertation, University of Cambridge).

References

- Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. Pantheon.
- Gullberg, M. (2006). Some reasons for studying gesture and second language acquisition. *International Review of Applied Linguistics in Language Teaching*, 44(2), 103–124.
- Gurrapu, V., Singh, A., & Sharma, P. (2025). Legal text summarization using abstractive models. *Proceedings of LREC-COLING 2025*.
- He, Y., Hu, T., Yu, J., Li, C., Weng, J., Gao, Z., & Wang, W. (2023). A survey on uncertainty quantification methods for deep learning. *arXiv preprint arXiv:2310.17321*.
- Holmes, J. (1995). *Women, men and politeness*. Longman.
- Hoetjes, M. (2024). Gesture and bilingual speech: Disfluency as a multimodal phenomenon. *Journal of Multimodal Communication*, 14(2), 221–239.
- Huang, Q. (2024). Modality matching for efficient and precise text interpretation: Experimentation with large language models. *Preprint*.
- Huang, X., Shen, Y., & Wang, L. (2024). Pragmatic inference in large language models. *Proceedings of ACL 2024*.
- Jones, C., Trott, S., & Bergen, B. (2024). Do multimodal large language models and humans ground language similarly? *Transactions of the Association for Computational Linguistics*, 12, 1456–1475.
- Kallmeyer, L. (2022). Disfluencies in multimodal communication. *Journal of Multimodal Communication*, 13(3), 199–212.
- Khujaniyazova, D. (2023). Institutional discourse in courtroom interaction. *Journal of Legal Communication*, 29(2), 145–160.
- König, P., Müller, L., & Stein, R. (2024). Governance of AI in legal systems. *AI & Society*.
- Kore, A., Mishra, R., & Singh, S. (2020). Extractive summarization using LexRank and TextRank. *Proceedings of COLING 2020*.
- Kryscinski, W., McCann, B., Xiong, C., & Socher, R. (2020). Evaluating the factual consistency of abstractive summarization. *Proceedings of EMNLP 2020*.
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.
- Lakoff, R. (1973). Language and woman's place. *Language in Society*, 2(1), 45–80.

References

- Leláková, E., & Praženicová, D. (2024). Exploring hedging in spoken discourse: Insights from corpus analysis. *Arab World English Journal*, 15(3), 270–282.
<https://doi.org/10.24093/awej/vol15no3.16>
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *Proceedings of ACL 2020*.
<https://doi.org/10.18653/v1/2020.acl-main.703>
- Li, J., Yuan, Y., & Zhang, Z. (2025). Legal domain-specific summarization with Longformer. *Proceedings of NAACL 2025*.
- Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. *Proceedings of the ACL Workshop on Text Summarization Branches Out*.
- Liu, X., Chen, T., Da, L., Chen, C., Lin, Z., & Wei, H. (2025). Uncertainty quantification and confidence calibration in large language models: A survey. *Proceedings of KDD '25* (pp. 1–11). ACM. <https://doi.org/10.1145/3711896.3736569>
- Loy, J., Rohde, H., & Corley, M. (2018). Cues to lying may be deceptive: Speaker and listener behavioural sensitivity to incongruence in speech content and confidence. *Acta Psychologica*, 191, 42–52.
- Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. *Proceedings of ACL 2020*.
- Mujtaba, D., Mahapatra, N. R., Arney, M., Yaruss, J. S., Gerlach-Houck, H., Herring, C., & Bin, J. (2024). Lost in transcription: Identifying and quantifying the accuracy biases of ASR systems against disfluent speech. *Proceedings of NAACL 2024* (pp. 4795–4809).
- Nan, L., Durmus, E., He, H., & Hashimoto, T. (2021). Entity-level factual consistency in abstractive summarization. *Proceedings of NAACL 2021*.
- Nam, J., & Jang, M. (2024). A survey on multimodal bidirectional machine learning translation of image and natural language processing. *Expert Systems with Applications*, 253, 124349.
- Nguyen, H., Wu, Z., & Lee, J. (2023). Sociolinguistic adaptation in neural models. *Proceedings of EMNLP 2023*.
- Nursafira, [Initials]. (2020). [Full reference needed — placeholder].
- Patel, R., & Joshi, A. (2024). Pragmatic inference in multilingual contexts. *Journal of Pragmatics*, 205, 55–70.

References

- Prince, E. F., Frader, J., & Bosk, C. (1982). On hedging in physician–physician discourse. In *Linguistics and the professions* (pp. 83–97). Ablex.
- Pu, J., Zhao, W., & Li, M. (2023). Automatic metrics for factual consistency. *Proceedings of EMNLP 2023*.
- Rahman, [Initials]. (2023). [Full reference needed — placeholder].
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140), 1–67.
- Safdar, M., Mahmood, U., & Rehman, A. (2024). Lexical hedges in female speech: An analysis of a contemporary talk show. *Pakistan Languages and Humanities Review*, 8(3), 611–617.
- Shen, L., Wu, Y., & Li, X. (2023). Bias and hedging in LLM summarization. *Proceedings of ACL 2023*.
- Shriberg, E. (2001). To “errrr” is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31(1), 153–169.
<https://doi.org/10.1017/S0025100301001128>
- Singh, A. (2024). Hybrid extractive–abstractive methods for legal document summarization. *Proceedings of LREC 2024*.
- Teleki, M., Deguchi, A., Samani, L., & Radulovic, F. (2024). Quantifying the impact of disfluency on spoken content summarization. *Proceedings of LREC-COLING 2024* (pp. 10877–10888). ACL.
- Tanneru, S. H. V., Agarwal, C., & Lakkaraju, H. (2024). Quantifying uncertainty in natural language explanations of large language models. *Proceedings of ICML 2024 (Vol. 238, pp. 34116–34134)*. PMLR.
- Varttala, T. (2001). *Hedging in scientifically oriented English*. Tampere University Press.
- Wagner, D., Bayerl, S. P., Baumann, I., Riedhammer, K., Nöth, E., & Bocklet, T. (2024). Large language models for dysfluency detection in stuttered speech. *Proceedings of Interspeech 2024*. International Speech Communication Association.
- Wu, J., Zhang, M., & Liu, Y. (2024). Sociolinguistic variation in neural summarizers. *Proceedings of ACL 2024*.
- Ye, F., Yang, M., Pang, J., Wang, L., Wong, D. F., Yilmaz, E., ... Tu, Z. (2024). Benchmarking LLMs via uncertainty quantification. *NeurIPS 2024*.

References

- Yona, G., Aharoni, R., & Geva, M. (2024). Can large language models faithfully express their intrinsic uncertainty in words? *Proceedings of EMNLP 2024* (pp. 7752–7764).
- Zhang, J., Zhao, Y., Saleh, M., & Liu, P. J. (2020). PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization. *Proceedings of ICML 2020*.
- Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). BERTScore: Evaluating text generation with BERT. *arXiv preprint arXiv:1904.09675*.
- Zhang, X., Li, M., & Wu, J. (2024). Conditional language learning with context. *Proceedings of ICML 2024*. ACM.
- Zheng, H., Xu, J., Li, Y., & Zhang, C. (2023). Detecting linguistic bias in legal corpora. *Proceedings of ICAIL 2023*.
- Zhong, H., & Litman, D. (2020). Automated fact extraction for legal case summarization. *Proceedings of LREC 2020*.
- Zhong, H., et al. (2020). JEC-QA: A legal-domain question answering dataset. *Proceedings of COLING 2020*.
- Zhong, H., Wang, C., & Li, Y. (2021). DialogSum: Dialogue summarization dataset. *Proceedings of ACL 2021*.