

Kyle Gorman

Curriculum vitae

December 9, 2020

852 Union St., Brooklyn, NY 11215. Tel. (267) 255-5669.

Education

Postdoctoral fellow, Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, 2013.

Ph.D., Linguistics, University of Pennsylvania, Philadelphia, PA, 2013.

B.A., Linguistics, University of Illinois at Urbana-Champaign, Urbana, IL, 2006.

Positions

Assistant professor, Graduate Program in Linguistics, Graduate Center, City University of New York, New York, NY, 2018–present.

Director, master's program in computational linguistics, Graduate Program in Linguistics, Graduate Center, City University of New York, New York, NY, 2018–present.

Software engineer, Google, Inc., New York, NY, 2015–present.

Assistant professor, Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, 2014–2015.

Publications

Books

In press. K. Gorman & R. Sproat. *Finite-State Text Processing*. Morgan & Claypool.

Refereed journal articles

2019. H. Zhang, R. Sproat, A. Ng, F. Stahlberg, X. Peng, K. Gorman, & B. Roark. Neural models of text normalization for speech applications. *Computational Linguistics* 45(2): 293–337.

2017. H. MacFarlane, K. Gorman, R. Ingham, A. Presmanes Hill, K. Papadakis, G. Kiss, & J. van Santen. Mazes in children with autism spectrum disorders or language impairment. *PLOS ONE* 12(3): e0173936.

2016. G. Fergadiotis, K. Gorman, & S. Bedrick. Algorithmic classification of five characteristic types of paraphasias. *American Journal of Speech-Language Pathology* 25(4S): S776–S787.
2016. K. Gorman & R. Sproat. Minimally-supervised number normalization. *Transactions of the Association for Computational Linguistics* 4: 507–519.
2016. K. Gorman, L. Olson, A. Presmanes Hill, R. Lunsford, P. Heeman, & J. van Santen. *Uh* and *um* in children with autism spectrum disorders or language impairment. *Autism Research* 9(8): 854–865.
2015. A. Presmanes Hill, J. van Santen, K. Gorman, B. Langhorst, & E. Fombonne. Memory in language-impaired children with and without autism. *Journal of Neurodevelopmental Disorders* 7: 19.
2013. L. Hinrichs, A. Bohmann, & K. Gorman. Real-time trends in the Texas English vowel system: F2 trajectory in GOOSE as an index of a variety’s ongoing delocalization. *Rice Working Papers in Linguistics* 4.
2011. K. Gorman, J. Howell, & M. Wagner. Prosodylab-Aligner: A tool for forced alignment of laboratory speech. *Journal of the Canadian Acoustical Association* 39(3): 192–193.
2011. J. Fruehwald & K. Gorman. Cross-derivational feeding is epiphenomenal. *Studies in the Linguistic Sciences* 2011: 36–50.

Refereed chapters

2019. K. Gorman & C. Yang. When nobody wins. In F. Rainer, F. Gardani, W. U. Dressler & H. C. Luschützky (ed.), *Competition in inflection and word formation*, 169–193. Dordrecht: Springer.
2013. K. Gorman & D. E. Johnson. Quantitative analysis. In R. Bayley, R. Cameron & C. Lucas (ed.), *The Oxford handbook of sociolinguistics*, 214–240. Oxford: Oxford University Press.

Refereed conference proceedings

2020. A. Waller & K. Gorman. Detecting objectifying language in online professor reviews. In *Proceedings of the Sixth Workshop on Noisy User-generated Text*, 171–180. **Best Paper Award.**
2020. P. Szymański & K. Gorman. Is the best better? Bayesian statistical comparison for natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2203–2212.
2020. K. Gorman, L. F.E. Ashby, A. Goyzueta, A. D. McCarthy, S. Wu, and D. You. The SIGMORPHON 2020 shared task on multilingual grapheme-to-phoneme conversion. In *17th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, 40–50.
2020. J. L. Lee, L. F.E. Ashby, M. E. Garza, Y. Lee-Sikka, S. Miller, A. Wong, A. D. McCarthy, and K. Gorman. Massively multilingual pronunciation mining with WikiPron. in *Proceedings of the 12th Language Resources and Evaluation Conference*, 4223–4228.
2020. A. D. McCarthy, C. Kirov, M. Grella, A. Nidhi, P. Xia, K. Gorman, E. Vylomova, S. J. Mielke, G. Nicolai, M. Silfverberg, T. Arkhangelskij, N. Krizhanovskiy, A. Krizhanovskiy, E. Klyachko, A. Sorokin, J. Mansfield, V. Ernštreits, Y. Pinter, C. L. Jacobs, R. Cotterell, M. Hulden, and D. Yarowsky. UniMorph 3.0: universal morphology. in *Proceedings of the 12th Language Resources and Evaluation Conference*, 3922–3931.
2019. K. Gorman, A. D. McCarthy, R. Cotterell, E. Vyolmova, M. Silfverberg, and M. Markowska. Weird inflects but OK: making sense of morphological generation errors. In *Proceedings of the 23rd Conference on Computational Natural Language Learning*, 140–151.
2019. S. Ritchie, R. Sproat, K. Gorman, D. van Esch, C. Schallhart, N. Bampounis, B. Brard, J. F. Mortensen, M. Holt, and E. Mahon. Unified verbalization for speech recognition & synthesis across languages. In *Proceedings of INTERSPEECH*, 3530–3534.

2019. K. Gorman & S. Bedrick. We need to talk about standard splits. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2786–2791. **Outstanding Paper Award.**
2019. S.J. Mielke, R. Cotterell, K. Gorman, B. Roark & J. Eisner. What kind of language is hard to language-model? In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 4975–4989.
2018. K. Gorman, G. Mazovetskiy, & V. Nikolaev. Improving homograph disambiguation with machine learning. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, 1349–1352.
2017. A. Ng., K. Gorman, & R. Sproat. Minimally supervised written-to-spoken text normalization. In *Proceedings of ASRU*, 665–670.
2017. J. Adams, S. Bedrick, G. Fergadiotis, K. Gorman, & J. van Santen. Target word prediction and paraphasia classification in spoken discourse. In *Proceedings of the ACL Workshop on Biomedical Natural Language Processing*, 1–8.
2016. K. Gorman. Pynini: A Python library for weighted finite-state grammar compilation. In *Proceedings of the ACL Workshop on Statistical NLP and Weighted Automata*, 75–80.
2015. K. Gorman, S. Bedrick, G. Kiss, M. Mohammed, R. Ingham, K. Papadakis, & J. van Santen. Automated morphological analysis of clinical language samples. In *Proceedings of ACL Workshop on Computational Linguistics and Clinical Psychology*, 108–116.
2014. M. Lehr, K. Gorman, & I. Shafran. Discriminative pronunciation model for dialectal speech recognition. In *Proceedings of INTERSPEECH*, 1458–1462.
2012. C. Lignos & K. Gorman. Revisiting frequency and storage in morphological processing. In *Proceedings of the 48th annual meeting of the Chicago Linguistic Society*, 447–461.
2012. K. Gorman. Rhotacism in Latin. In *Proceedings of the 48th annual meeting of the Chicago Linguistic Society*, 279–293.
2011. K. Gorman. A program for phonotactic theory. In *Proceedings of the 47th annual meeting of the Chicago Linguistic Society*, 79–93.
2010. K. Gorman. The consequences of multicollinearity among socioeconomic predictors of negative concord in Philadelphia. In M. Lerner, ed., *U. Penn Working Papers in Linguistics 16.2: Selected papers from NAWAV 38*, 66–75.
2007. C. Lai, K. Gorman, J. Yuan, & M. Liberman. Perception of disfluency: Language differences and listener bias. In *Proceedings of INTERSPEECH*, 2345–2348.

Manuscripts in preparation

- K. Gorman. CityLex: a free English lexical database.
- K. Gorman. A* shortest-path decoding for non-path semirings.
- K. Gorman & P. Bryan. Sex differences in English irregular verb acquisition.

Technical reports

- K. Gorman. Hierarchical regression for language research. Institute for Research in Cognitive Science Technical Report 09–02, 2009.

Grants & awards

2019. Association for Computational Linguistics, Outstanding Paper Award (with S. Bedrick); \$1,000.

2014–2015. NIH R21DC014099: Co-construction of lexica in primary progressive aphasia (Co-investigator; PI: M. Fried-Oken); \$429,066.

2011–present. NIH R01DC012033, Computational characterization of language use in autism (Co-investigator; PI: J. van Santen).

Patents

2019. R. Sproat, K. Wu, & K. Gorman. Semiotic class normalization. U.S. Patent 9,852,123.

Advising

Supervised MA theses

2020. Emily Campbell. Does the word "chien" bark? Representation learning in neural machine translation encoders.

2020. Alan Wong. Mitigating gender bias in neural machine translation using counterfactual data.

2020. Magdalena Markowska. Tones in Shupamem reduplication.

2020. Alina Korovatskaya. Genderlects in social media.

2020. Michael C. Stern. Testing the perceptual magnet effect in monolinguals and bilinguals. (Winner of the linguistics annual master's thesis prize.)

2020. Anthony J. Vicario. Phonologically-informed speech coding for automatic speech recognition-based foreign language pronunciation training.

2020. Angie Waller. Ghost peppers: using ensemble models to detect professor attractiveness commentary on RateMyProfessors.com. (Winner of the computational linguistics annual master's thesis prize.)

2019. Sara Morini. Demographic factors as domains for adaptation in linguistic preprocessing.

Invited talks

2020. Anatomy of an analogy. Department of Linguistics, Stony Brook University, November 6.

2020. Abstractness, arbitrariness, and productivity in Polish declension. Graduate Program in Linguistics, Graduate Center, City University of New York, October 8.

2020. Finite-state text processing. North American Summer School for Logic, Language, and Information (NASSLLI), Brandeis University, Boston, July 11. (Cancelled due to pandemic.)

2020. Asymmetries in Latin glide formation. North American Phonology Conference (NAPhC), Concordia University, Montreal, May 8–9. (Cancelled due to pandemic.)

2020. Weird inflects but OK: making sense of morphological generation errors. Linguistics Department, University of California, San Diego, January 30.

2020. A tutorial on finite-state text processing. Society for Computation in Linguistics, New Orleans, January 2.

2019. Massively multilingual resources for speech and language technology. Department of Linguistics and Communication Disorders, Queens College, December 12.

2019. Weird inflects but OK: making sense of morphological generation errors. Northeast Computational Phonology Circle (NECPhon), Rutgers University, November 16.

2019. A tutorial on finite-state text processing. Linguistic Data Consortium Institute, November 5.

2019. We need to talk about standard splits. Fondazione Bruno Kessler, July 25.

2019. Two little pieces on statistical system evaluation and error analysis. Sapienza University of Rome, July 24.

2019. Internationalization challenges for speech and language technologies. Documentation Division, United Nations, February 19.

2019. Statistically informed reproduction analysis. Linguistics Department, Swarthmore College, February 12.

2019. Statistically informed reproduction analysis. Institute for Research in Cognitive Science, University of Pennsylvania, February 11.

2018. Grammar engineering for speech and language processing. Center for Data Science, New York University, November 29.

2018. Grammar engineering for speech and language processing. Linguistics Department, McGill University, November 14.

2018. Finite-state grammar development in Pynini. Linguistics Department, McGill University, November 12.

2018. Grammar engineering for speech and language processing. Program in Linguistics, Princeton University, September 26.

2018. Minimally supervised number normalization. RIKEN Center for Advanced Intelligence, May 16.

2018. Minimally supervised number normalization. Information Processing Society of Japan, Tokyo, May 12.

2018. Minimally supervised number normalization. University of California, Irvine, May 4.

2018. Computational approaches to language. Graduate Program in Linguistics, Graduate Center, City University of New York, March 23.

2017. Minimally supervised number normalization. Institute for Advanced Computational Science, Stony Brook University, November 29.

2016. Processing informal text. Department of English, University of Texas at Austin, November 28.

2014. Forced alignment with Prosodylab-Aligner. Experimental and Theoretical Advances in Prosody, McGill University, May 10.

2013. Computational characterizations to language in children with autism spectrum disorders. University of Connecticut, Storrs, October 16.

2013. What phonotactics might not be. Department of Linguistics, University of Oregon, April 2.

2013. What phonotactics might not be. Department of Linguistics, Reed College, January 31.

2012. Words vs. rules in irregular acquisition. Center for Spoken Language Understanding, Oregon Health & Science University, June 13.

2011. Forced alignment with Prosodylab-Aligner. Chicago Linguistics Society, University of Chicago, October 14.

2011. Quantitative analysis for linguists. Hunter College, September 9.

2011. What phonotactics might not be. Department of Linguistics, New York University, March 4.

2011. What phonotactics might not be. Department of Linguistics, University of California, Santa Cruz, February 10.

2011. Latin rhotacism for real. Department of Linguistics, Concordia University, February 4.

Teaching & service

Graduate teaching

Methods in Computational Linguistics I, The Graduate Center, City University of New York.

Methods in Computational Linguistics II, The Graduate Center, City University of New York.

Language Technology, The Graduate Center, City University of New York.

Statistics for Linguistics Research, The Graduate Center, City University of New York.

Seminar in Writing Systems, The Graduate Center, City University of New York.

Natural Language Processing, Oregon Health & Science University.

Research Programming, Oregon Health & Science University.

Internal service

2019–present. Director, Speaker Series in Computational Linguistics.

2019–present. Member, Executive Committee, Graduate Program in Linguistics.

2019–present. Member, Admissions Committee, Graduate Program in Linguistics.

2019–present. Member, Graduate Council, the Graduate Center.

External service

2020. Area Chair, Conference on Natural Language Learning, Association for Computational Linguistics (ACL) Special Interest Group on Natural Language Learning (SIGNLL).

2019–present. Executive Committee, Association for Computational Linguistics (ACL) Special Interest Group on Computational Morphology and Phonology (SIGMORPHON).

2018. Chair of program committee for the Language Resources & Evaluation (LREC) industry track.

Reviewer for ACL, EAACL, EMNLP, ICASSP, INTERSPEECH, LREC, the National Science Foundation, the Natural Sciences and Engineering Research Council of Canada, Oxford University Press, *Cognition*, *Cognitive Science*, *Computational Linguistics*, *Journal of Linguistics*, *Journal of Autism & Developmental Disorders*, *Language*, *Language Variation & Change*, *Lingua*, *Phonology*, and *PLOS ONE*,

Professional affiliations

Linguistics Society of America

Association for Computational Linguistics

Institute of Electrical and Electronics Engineers