

Gus Hahn-Powell

University of Arizona
Department of Linguistics
Douglass Building Room 228
Tucson, AZ 85721

✉ gus@parsertongue.org
🌐 <https://parsertongue.org>
👤 github.com/myedibleenso
📞 0000-0002-7957-3586

Employment

- 2022-present **Amazon**
Applied Scientist, Amazon Product Graph (Personalization)
- 2019-present **University of Arizona**
Assistant Professor, Department of Linguistics
Director, Human Language Technology (HLT) Online MS Program
Director, Natural Language Processing (NLP) Graduate Certificate Program
Faculty, Cognitive Science Graduate Interdisciplinary Program
Faculty, Computational Social Science Graduate Certificate Program
- 2017-present **LUM AI**
Co-founder and Solutions Architect

Education

- 2013-2018 **Ph.D.** in Computational Linguistics, *University of Arizona*
Dissertation: Machine Reading for Scientific Discovery
- 2012-2014 **M.S.** Human Language Technology, *University of Arizona*
- 2008-2010 **M.A.** Applied Linguistics, *University of Alabama*
- 2004-2008 **B.A.** Japanese, *University of Alabama*
- 2006-2007 Study abroad in Kyoto, Japan (*Ritsumeikan University*)

Publications (peer-reviewed)

- 2022 Enrique Noriega-Atala, Robert Vacareanu, Gus Hahn-Powell, and Marco A. Valenzuela-Escárcega (Oct. 2022). “Neural-Guided Program Synthesis of Information Extraction Rules Using Self-Supervision.” In: *Proceedings of the First Workshop on Pattern-based Approaches to NLP in the Age of Deep Learning*. Gyeongju, Republic of Korea: International Conference on Computational Linguistics, pp. 85–93. URL: <https://aclanthology.org/2022.pandl-1.10>
- Arie Sutiono and Gus Hahn-Powell (Oct. 2022). “Syntax-driven Data Augmentation for Named Entity Recognition.” In: *Proceedings of the First Workshop on Pattern-based Approaches to NLP in the Age of Deep Learning*. Gyeongju, Republic of Korea: International Conference on Computational Linguistics, pp. 56–60. URL: <https://aclanthology.org/2022.pandl-1.7>

- Robert Vacareanu, George C.G. Barbosa, Enrique Noriega-Atala, Gus Hahn-Powell, Rebecca Sharp, Marco A. Valenzuela-Escárcega, and Mihai Surdeanu (July 2022). “A Human-machine Interface for Few-shot Rule Synthesis for Information Extraction.” In: *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations*. Hybrid: Seattle, Washington + Online: Association for Computational Linguistics, pp. 64–70. URL: <https://aclanthology.org/2022.naacl-demo.8>
- Robert Vacareanu, Marco A. Valenzuela-Escárcega, George Barbosa, Rebecca Sharp, Gus Hahn-Powell, and Mihai Surdeanu (2022). “From Examples to Rules: Neural Guided Rule Synthesis for Information Extraction.” In: *Proceedings of the 13th Language Resources and Evaluation Conference (LREC)*. URL: <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.665.pdf>
- 2021 Elsayed Issa, Mohammed AlShakhori, Reda Al-Bahrani, and Gus Hahn-Powell (2021). “Country-level Arabic dialect identification using RNNs with and without linguistic features.” In: *Proceedings of the Sixth Arabic Natural Language Processing Workshop*. Kiev, Ukraine (Online): Association for Computational Linguistics. URL: <https://aclanthology.org/2021.wanlp-1.32>
- 2020 Marco A. Valenzuela-Escárcega, Gus Hahn-Powell, and Dane Bell (May 2020). “Odinson: A Fast Rule-based Information Extraction Framework.” In: *Proceedings of The 12th Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association, pp. 2183–2191. URL: <https://aclweb.org/anthology/2020.lrec-1.267>
- Zheng Tang, Gus Hahn-Powell, and Mihai Surdeanu (July 2020). “Exploring Interpretability in Event Extraction: Multitask Learning of a Neural Event Classifier and an Explanation Decoder.” In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*. Online: Association for Computational Linguistics, pp. 169–175. URL: <https://aclweb.org/anthology/2020.acl-srw.23>
- 2019 Robert Poole, Andrew Gnan, and Gus Hahn-Powell (2019). “Epistemic stance and the construction of knowledge in science writing: A diachronic corpus study.” In: *Journal of English for Academic Purposes* 42, p. 100784. ISSN: 1475-1585. DOI: 10.1016/j.jeap.2019.100784. URL: <https://parsertongue.org/preprints/construction-of-knowledge-in-science/>
- George C. G. Barbosa, Zechy Wong, Gus Hahn-Powell, Dane Bell, Rebecca Sharp, Marco A. Valenzuela-Escárcega, and Mihai Surdeanu (June 2019). “Enabling Search and Collaborative Assembly of Causal Interactions Extracted from Multilingual and Multi-domain Free Text.” In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*. ACL. Minneapolis, Minnesota: Association for Computational Linguistics, pp. 12–17. DOI: 10.18653/v1/N19-4003. URL: <https://aclweb.org/anthology/N19-4003/>

- 2018 Marco A. Valenzuela-Escárcega, Özgün Babur, Gus Hahn-Powell, Dane Bell, Thomas Hicks, Enrique Noriega-Atala, Xia Wang, Mihai Surdeanu, Emek Demir, and Clayton T. Morrison (2018). “Large-scale Automated Machine Reading Discovers New Cancer Driving Mechanisms.” In: *Database: The Journal of Biological Databases and Curation*. DOI: 10.1093/database/bay098
- Fan Luo, Marco A. Valenzuela-Escárcega, Gus Hahn-Powell, and Mihai Surdeanu (2018). “Scientific Discovery as Link Prediction in Influence and Citation Graphs.” In: *Proceedings of the Twelfth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-12)* (New Orleans, Louisiana, USA). Association for Computational Linguistics, pp. 1–6. DOI: 10.18653/v1/W18-1701
- Angus Graeme Forbes, Kristine Lee, Gus Hahn-Powell, Marco Antonio Valenzuela-Escárcega, and Mihai Surdeanu (2018). “Text Annotation Graphs: Annotating Complex Natural Language Phenomena.” In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)* (Miyazaki, Japan). European Language Resources Association (ELRA). arXiv: 1711.00529 [cs.CL]. URL: <https://aclweb.org/anthology/L18-1169/>
- Heather Lent, Gus Hahn-Powell, Asher Haug-Baltzell, Sean Davey, Mihai Surdeanu, and Eric Lyons (2018). “Science Citation Knowledge Extractor.” In: *Frontiers in Research Metrics and Analytics*. Ed. by Neil Smalheiser. DOI: 10.3389/frma.2018.00035
- 2017 Gus Hahn-Powell, Marco A. Valenzuela-Escárcega, and Mihai Surdeanu (2017). “Swanson linking revisited: Accelerating literature-based discovery across domains using a conceptual influence graph.” In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics: Software Demonstrations*. ACL, pp. 103–108. DOI: 10.18653/v1/P17-4018
- Marco A Valenzuela-Escárcega, Ozgün Babur, Gus Hahn-Powell, Dane Bell, Thomas Hicks, Enrique Noriega-Atala, Xia Wang, Mihai Surdeanu, Emek Demir, and Clayton T Morrison (2017). “Large-scale automated reading with Reach discovers new cancer driving mechanisms.” In: *Proceedings of the Sixth BioCreative Challenge Evaluation Workshop*, pp. 201–203. URL: http://www.biocreative.org/media/store/files/2018/general_3.pdf
- 2016 Gus Hahn-Powell, Dane Bell, Marco A. Valenzuela-Escárcega, and Mihai Surdeanu (2016). “This before That: Causal Precedence in the Biomedical Domain.” In: *Proceedings of the 2016 Workshop on Biomedical Natural Language Processing* (Humboldt University of Berlin). Association for Computational Linguistics, pp. 146–155. DOI: 10.18653/v1/W16-2920. arXiv: 1606.08089 [cs.CL]
- Dane Bell, Gus Hahn-Powell, Marco A. Valenzuela-Escárcega, and Mihai Surdeanu (2016). “An investigation of coreference phenomena in the biomedical domain.” In: *Proceedings of the 10th International Conference on Language Resources and Evaluation* (Portorož, Slovenia). LREC. arXiv: 1603.03758 [cs.CL]. URL: <https://aclweb.org/anthology/L16-1027/>

Marco A. Valenzuela-Escárcega, Gus Hahn-Powell, Dane Bell, and Mihai Surdeanu (2016). “SnapToGrid: From Statistical to Interpretable Models for Biomedical Information Extraction.” In: *Proceedings of the 15th Workshop on Biomedical Natural Language Processing* (Humboldt University of Berlin). Association for Computational Linguistics, pp. 56–65. DOI: 10.18653/v1/W16-2907. arXiv: 1606.09604 [cs.CL]

Marco A. Valenzuela-Escárcega, Gus Hahn-Powell, and Mihai Surdeanu (2016). “Odin’s Runes: A Rule Language for Information Extraction.” In: *Proceedings of the 10th International Conference on Language Resources and Evaluation* (Portorož, Slovenia). LREC. URL: <https://aclweb.org/anthology/L16-1050/>

2015 Daniel Fried, Peter Jansen, Gus Hahn-Powell, Mihai Surdeanu, and Peter Clark (2015). “Higher-order Lexical Semantic Models for Non-factoid Answer Reranking.” In: *Transactions of the Association for Computational Linguistics* 3, pp. 197–210. ISSN: 2307-387X. URL: <https://aclweb.org/anthology/Q15-1015/>

Marco A. Valenzuela-Escárcega, Gus Hahn-Powell, Thomas Hicks, and Mihai Surdeanu (2015). “A Domain-independent Rule-based Framework for Event Extraction.” In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing: Software Demonstrations* (Beijing, China). ACL-IJCNLP 2015, pp. 127–132. DOI: 10.3115/v1/P15-4022

Presentations (no proceedings)

2015 Gus Hahn-Powell, Benjamin Martin, and Diana Archangeli (Dec. 2015). “A method for automatically detecting problematic tongue traces.” In: *Proceedings of Ultrafest VII* (University of Hong Kong). Ultrafest VII. URL: <https://parsertongue.org/presentations/2015/ultrafest-tongue-traces/>

2014 Gus Hahn-Powell and Diana Archangeli (Oct. 2014a). “AutoTrace: An automatic system for tracing tongue contours.” In: *Proceedings of the 168th Meeting of Acoustical Society of America* (Indianapolis, Indiana). Vol. 136. 4. ASA, pp. 2104–2104. URL: <https://parsertongue.org/presentations/2014/asa-at/>

Gus Hahn-Powell and Diana Archangeli (Oct. 2014b). “Testing AutoTrace.” In: *Proceedings of the 168th Meeting of Acoustical Society of America* (Indianapolis, Indiana). Vol. 136. 4. ASA, pp. 2082–2082. URL: <https://parsertongue.org/presentations/2014/asa-testing-at/>

2013 Diana Archangeli, Mohsen Mahdavi, David Ellison, Gus Hahn-Powell, Rolando Coto, Jeff Berry, and Paul Boersma (Nov. 2013). “UltraPraat Software & database for simultaneous acoustic and articulatory analysis.” In: *Proceedings of Ultrafest VI* (Queen Margaret University). Ultrafest VI

Jae-Hyun Sung, Jeff Berry, Marissa Cooper, Gus Hahn-Powell, and Diana Archangeli (Nov. 2013). “Testing AutoTrace: A Machine-learning Approach to Automated Tongue Contour Data Extraction.” In: *Proceedings of Ultrafest VI* (Queen Margaret University). Ultrafest VI. URL: <https://parsertongue.org/presentations/2013/ultrafest-at/>

Colin R. Dawson, Luca Del Pero, Clayton T. Morrison, Mihai Surdeanu, Gus Hahn-Powell, Zachary Chapman, and Kobus Barnard (Apr. 2013). “Bayesian modeling of scenes and captions.” In: *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; Workshop on Vision and Language* (Atlanta, Georgia). (WVL)NAACL-HLT. URL: http://nlp.cs.illinois.edu/WVL13/slides/Dawson_WVL13.pdf

2010 Elliot Patton, Gus Hahn-Powell, and Robert Nelson (Apr. 2010). “The ‘Worthy of Attention’ Collostruction: Frequency, synonymy, and learnability.” In: *Southeastern Conference on Linguistics* (University of Mississippi). SECOL LXII. URL: <https://parsertongue.org/presentations/2010/secol-worthy-of-attention-collostruction/>

Forthcoming

Neil R. Smalheiser, Gus Hahn-Powell, Dimitar Hristovski, and Yakub Sebastian (2023). “From Knowledge Discovery to Knowledge Creation: How can Literature-based Discovery Accelerate Progress in Science?” In: *OECD iLibrary*. Ed. by Organisation for Economic Cooperation and Development (OECD)

Invited Talks

2021 Gus Hahn-Powell (Oct. 2021). *Human Language Technology Webinar*. UNAM. National Autonomous University of Mexico (UNAM). URL: https://twitter.com/UNAM_Tucson/status/1453402189607620613

2020 Gus Hahn-Powell (May 2020a). *Community-guided Hypothesis Generation*. PAKDD. First International Workshop on Literature-Based Discovery (LBD 2020). URL: <https://www.pakdd2020.org/workshops.html>

Gus Hahn-Powell (Apr. 2020b). *Generating scientific hypotheses through machine reading*. UACOGSCI. University of Arizona’s Cognitive Science Colloquium series. URL: <https://cogsci.arizona.edu/content/cognitive-science-colloquium>

2019 Gus Hahn-Powell and Dane Bell (Oct. 2019). *Bridging Non-interacting Research Communities Through Machine-guided Discovery Synthesis*. To be presented at the INFORMS 2019 special session on Machine Reading and Comprehension for Science-Practice Knowledge Synthesis. Seattle, WA. INFORMS. URL: <https://www.abstractsonline.com/pp8/#!/6818/presentation/6580>

Patents

Pending

- 2018 | Mihai Surdeanu, Marco Valenzuela-Escárcega, Gus Hahn-Powell, Dane Bell, Thomas Hicks, Enrique Noriega, and Clayton Morrison (Sept. 2018). “Methods for extracting and assessing information from literature documents.” Patent US 20180260474A1 (US). URL: <https://patentimages.storage.googleapis.com/a6/82/53/42c8519df48fe8/US20180260474A1.pdf>

Publications (not peer-reviewed)

Manuals

- 2015 | Marco Antonio Valenzuela-Escárcega, Gus Hahn-Powell, and Mihai Surdeanu (2015). *Description of the Odin Event Extraction Framework and Rule Language*. v1. arXiv: 1509.07513 [cs.CL]

Grants and Awards

Grants

- 2021 | *ADHS-CDC COVID Disparities Initiative*. AGENCY: CDC & Arizona Department of Health Services. URL: <https://crh.arizona.edu/programs/covid-disparities-initiative>. ROLE: Co-I. AWARD: \$8M
- 2020 | *Democratizing machine reading for non-experts: Easy and interpretable methods to extract structured information from text*. AGENCY: NSF. URL: <https://mr4all.parsertongue.org/about>. ROLE: Co-PI. FEDERAL AWARD ID: 2006583. AWARD: \$499K
- Supply chain Quantification Using Imperfect Data (SQUID)*. AGENCY: DARPA. URL: <https://darpa.mil/program/logx>. ROLE: PI (**Phase I** subcontract through Raytheon BBN). AWARD: \$282K

Awards

- 2019 | **Best System Demonstration**. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*. URL: <https://influence.demos.lum.ai>

Research Interests

machine reading, literature-based discovery, knowledge assembly, lexical semantics, computational modeling of language