

# Alexandra Salem

---

## Education

2018–2021 **MS in Computer Science & Engineering**, Oregon Health & Science University, Portland, OR.

### Masters Final Project

Fine-tuned the machine learning based large language model BERT for automatic semantic similarity classification of responses to items in the picture-naming evaluation of aphasia known as the Philadelphia Naming Test (PNT). This work was published in the Journal of Speech Language & Hearing Research (see [https://doi.org/10.1044/2022\\_JSLHR-22-00277](https://doi.org/10.1044/2022_JSLHR-22-00277)).

### Relevant Coursework

Image Processing, Natural Language Processing, Machine Learning, Advanced Machine Learning, Deep Learning, Deep Learning II, Introduction to Database Management, Problem Solving with Large Clusters, Research Ethics for Computer Science & Electrical Engineering, Data Science Programming, Data Visualization, Probability & Statistics

2012–2016 **BA in Mathematics**, Reed College, Portland, OR.

### Relevant Coursework

Introduction to Computing, Data Structures and Algorithms, Single and Multivariable Calculus, General Physics I & II, Probability Theory, Real Analysis, Abstract and Linear Algebra, Combinatorics

---

## Work Experience

Nov 2016–Present **Research Staff, Department of Medical Informatics and Clinical Epidemiology**, Oregon Health & Science University, Portland, OR.

*Research Data Analyst II, Nov 2021– Present &*

*Senior Research Assistant, July 2021– Nov 2021.*

- Develop software applications in Python for automated assessment of clinical language data using machine learning.
- Independently design and execute scientific experiments for evaluating the efficacy of our assessment software utilizing advanced statistical methods.
- Analyze, visualize and summarize data findings using Python & R statistical software.
- Write scientific papers for journals and conferences, and present research at conferences to mixed audiences from clinical and computational fields.
- Contribute to writing grant applications and progress reports.
- Collaborate with experimental scientists from clinical and computational fields.

*Research Assistant II, Sept 2017– July 2021.*

- Designed data collection, data management, and data annotation protocols for audiovisual data derived from clinical contexts.
- Transcribed speech of children with autism using ELAN software.
- Contributed to scientific report writing and presented research at conferences to mixed audiences from clinical and computational fields.
- Conducted administrative duties such as coordinating team meetings and monitoring IRB compliance.

*Research Volunteer, Nov 2016– Sept 2017.*

- May 2022– July 2022 **Machine Learning Scientist Intern**, *Cambia Health Solutions*, Portland, OR.
- Retrained, evaluated, and deployed a machine learning classifier for filtering medical claims data. This model reduced member abrasion and claims workload.
  - Presented the model to technical and non-technical audiences.
- Feb 2017– Sept 2017 **Academic Specialist**, *North Avenue Education*, Portland, OR.
- Provided one-on-one tutoring sessions to students in Algebra, Geometry, Algebra 2, Probability & Statistics, and AP Calculus.
  - Planned individualized lessons for as many as ten students at a time.

---

## Skills

Python (experience with PyTorch, TensorFlow, Transformers, SciKit-Learn, Pandas, Numpy, PySpark etc), R (ggplot, tidyverse), SQL, Slurm, Git, HTML, CSS, JavaScript, L<sup>A</sup>T<sub>E</sub>X, MS Office, Audacity, ELAN, Mathematica

---

## Publications

Gale, R., Salem, A. C., Fergadiotis, G., & Bedrick, S. (2023). Mixed Orthographic/Phonemic Language Modeling: Beyond Orthographically Restricted Transformers (BORT). *8th Workshop on Representation Learning for NLP*. Association for Computational Linguistics. <https://aclanthology.org/2023.rep14nlp-1.18>

Salem, A. C., Gale, R., Fleegle, M., Fergadiotis, G., & Bedrick, S. (2023). Automating intended target identification for paraphasias in discourse using a large language model. *Journal of Speech, Language, and Hearing Research*. (Under review).

MacFarlane, H., Salem, A. C., Bedrick, S., Dolata, J. K., Wiedrick, J., Lawley, G. O., Finestack, L. H., Kover, S. T., Thurman, A. J., Abbeduto, L., & Fombonne, E. (2023). Consistency and reliability of automated language measures across expressive language samples in autism. *Autism Research*. <https://doi.org/10.1002/aur.2897>

Casilio, M., Fergadiotis, G., Salem, A. C., Gale, R., McKinney-Bock, K., & Bedrick, S. (2023). ParAlg: A paraphasia algorithm for multinomial classification of picture naming errors. *Journal of Speech, Language, and Hearing Research*. [https://doi.org/10.1044/2022\\_JSLHR-22-00255](https://doi.org/10.1044/2022_JSLHR-22-00255)

Salem, A. C., Gale, R., Casilio, M., Fleegle, M., Fergadiotis, G., & Bedrick, S. (2022). Refining semantic similarity of paraphasias using a contextual language model. *Journal of Speech, Language, and Hearing Research*. [https://doi.org/10.1044/2022\\_JSLHR-22-00277](https://doi.org/10.1044/2022_JSLHR-22-00277)

MacFarlane, H., Salem, A. C., Chen, L., Asgari, M., & Fombonne, E. (2022). Combining voice and language features improves automated autism detection. *Autism Research*, 15(7), 1288–1300. <https://doi.org/10.1002/aur.2733>

Salem, A. C., MacFarlane, H., Adams, J. R., Lawley, G. O., Dolata, J. K., Bedrick, S., & Fombonne, E. (2021). Evaluating atypical language in autism using automated language measures. *Scientific Reports*, 11(1), 10968. <https://doi.org/10.1038/s41598-021-90304-5>

Adams, J. R., Salem, A. C., MacFarlane, H., Ingham, R., Bedrick, S. D., Dolata, J. K., Hill, A. P., Fombonne, E., & van Santen, J. (2021). A pseudo-value approach to measuring lexico-semantic similarity in the language of children with or without autism. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2021.668344>

Fombonne, E., MacFarlane, H., & Salem, A. C. (2021). Epidemiological surveys of ASD: Advances and remaining challenges. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-021-05005-9>

MacFarlane, H., Salem, A. C., Lawley, G. O., Hill, A. P., Zuckerman, K., & Fombonne, E. (2022). Epidemiology of ASD. In *Textbook for Autism Spectrum Disorders* (2nd ed.). American Psychiatric Association Publishing.

---

## Presentations/Talks

Salem, A. C., Gale, R., Fleegle, M., Fergadiotis, G., & Bedrick, S. (2023, June). *Automating intended target identification for paraphasias in discourse using a large language model*. Poster presented by Alexandra Salem at: 52nd Clinical Aphasiology Conference, Atlantic City, NJ.

Salem, A. C., Gale, R., Fergadiotis, G., & Bedrick, S. (2021, October). *Improving Automatic Semantic Similarity Classification of the PNT*. Poster presented by Alexandra Salem at: Academy of Aphasia 59th Annual Meeting, Online.

Salem, A. C., Gale, R., & Bedrick, S. (2021, May). *Fine-tuning BERT for the semantic classifier of the PNT*. Short talk by Alexandra Salem at: Oregon Health & Science University Research Week, Portland, OR.

Salem, A. C., Gale, R., & Bedrick, S. (2021, March). *Fine-tuning BERT for the semantic classifier of the PNT*. Talk by Alexandra Salem at: Masters Final Project Department Talk, Oregon Health & Science University, Portland, OR.

Salem, A. C. (2018, August). *A sentimental analysis of Tolstoy*. Invited talk at: Portland R User Group Meeting, Portland, OR.

Adams, J. R., Salem, A. C., Hill, A. P., Bedrick, S., & van Santen, J. (2018, May). *Semantic similarity of conversational speech between children with and without ASD*. Poster presented by Joel Adams and Alexandra Salem at: Oregon Health & Science University Research Week, Portland, OR.

Adams, J. R., Salem, A. C., Hill, A. P., Bedrick, S., & van Santen, J. (2018, April 27). *Semantic similarity of conversational speech between children with and without ASD*. Short Talk by Joel Adams at: 5th Pacific Northwest Regional NLP Workshop: NW-NLP, 2018, Redmond, WA.

---