

Cristina M. Aggazzotti

✉ caggazz1@jh.edu 📧 caggazzotti 📧 c-aggazzotti

Education

Harvard University

PH.D. IN LINGUISTICS

M.A. IN LINGUISTICS

Cambridge, MA

2019

2017

University College London

M.Sc. IN COGNITIVE AND DECISION SCIENCES (Distinction)

London, UK

2013

University of Southern California

B.A. IN APPLIED AND COMPUTATIONAL MATHEMATICS

B.A. IN LINGUISTICS

Magna Cum Laude; Renaissance Scholar

Los Angeles, CA

2011

Research and Work Experience

Postdoctoral Fellow in Computer Science

JOHNS HOPKINS UNIVERSITY

Baltimore, MD

Aug. 2023 – Pres.

- Work toward robust and explainable authorship models by exploring model representations, biases, and performance
- Develop ground truth datasets for testing various kinds of authorship models

Academic Editor and Proofreader

FREELANCER

Remote

Feb. 2017 – Pres.

- Edit and proofread journal papers, books, job/grant applications, & Ph.D. dissertations for non-native English speakers

Postdoctoral Fellow in Linguistics

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Remote

May 2022 – Aug. 2023

- Developed and tested stylometric and neural authorship attribution models across genres, domains, and dataset sizes
- Identified forensic linguistic data needs and began developing standards for multidisciplinary data use

Visiting Research Scholar (Summer Camp for Applied Language Exploration)

HUMAN LANGUAGE TECHNOLOGY CENTER OF EXCELLENCE, JOHNS HOPKINS UNIVERSITY

Baltimore, MD

Jun. 2022 – Aug. 2022

- Experimented with schemas masking training data to explain neural author models' reliance on style vs. content
- Developed a linguistic and stylometric feature model for determining authorship of various text types

Research Contractor

INSTITUTE FOR LINGUISTIC EVIDENCE

Remote

2016 – 2018; Oct. 2021 – Apr. 2022

- Compared discriminant analysis (SPSS) and LSTM (Python) approaches to classifying suicide notes and last words
- Analyzed linguistic evidence across authors and text genres; researched statistical methods for presentation in court

Research Scientist and Customer Experience Engineer

POINTMAIL, INC. (STARTUP)

Cambridge, MA

Feb. 2018 – Jan. 2020

- Investigated ways to improve phrase- and sentence-level text prediction based on individual language usage and style
- Led bi-weekly team-wide research meetings and helped guide future research directions and product decisions
- Built rapport with customers, onboarded users, resolved customer issues, and elicited & evaluated customer feedback

Graduate Researcher

HARVARD UNIVERSITY

Cambridge, MA

Aug. 2015 – May 2019

- Created a unified analysis of reflexives and reciprocals, implemented in synchronous tree adjoining grammar (a syntax-semantics hierarchical tree pairing system), and provided a novel extension of the formalism to capture morphology
- Developed a syntactic and semantic annotation scheme for reflexives using the Corpus of Contemporary American English by creating a hybrid classification system based on previous reflexives research and attested data
- Mentored 7 undergrads in annotation schemes, machine learning of anaphor data, and testing my model using OCaml

Graduate Researcher

UNIVERSITY COLLEGE LONDON

London, UK

Sep. 2012 – Sep. 2013

- Created a formalized crime script of cash-in-transit robbery combining influence diagrams, causal models, Bayesian networks, and scripts that could ultimately be used by the police to identify effective prevention strategies

Teaching and Mentoring Experience

Research Mentor

Remote

POLYGENCE

Sep. 2022 – Sep. 2023

- Mentored high school students through a research project and advised on academics and possible career paths

Teaching Fellow

Cambridge, MA

HARVARD UNIVERSITY

Aug. 2016 – May 2018

- Assisted in teaching two intro linguistics classes (Distinction in Teaching Award for each) and an intro syntax class
- Created curricula for and taught as the primary instructor tutorials on computational linguistics and forensic linguistics

Tutor and Careers Class Assistant

Boston, MA

SUFFOLK COUNTY HOUSE OF CORRECTIONS; BOSTON COMM. CORRECTIONS CENTER

Feb. 2015 – May 2016

- Taught English to Spanish-, Portuguese-, and Mandarin-speaking inmates; tutored men and women on probation/parole for the GED and assisted in a careers class creating resumes and helping with job applications

Tutor and College Consultant

Newbury Park, CA

SELF-EMPLOYED

Sep. 2013 – Aug. 2014

- Tutored 30 students for the SAT, ACT, and all high school subjects; >300pt avg increase on SAT, 3pt avg increase on ACT
- Counseled 5 students through college selection, applications, resumes, and essays with successful admissions for all

Publications and Presentations

- Aggazzotti, Cristina, Nicholas Andrews, and Elizabeth Allyn Smith. **Can authorship attribution models distinguish speakers in speech transcripts?** (2023) cs.CL/2311.07564v2.
- Wang, Andrew, Cristina Aggazzotti, Rebecca Kotula, Rafael Rivera Soto, Marcus Bishop, and Nicholas Andrews. **Can authorship representation learning capture stylistic features?** Transactions of the Association for Computational Linguistics, 11 pp. 1416–1431 (2023).
- Aggazzotti, Cristina and Elizabeth Allyn Smith. **Disambiguating ‘ground truth’ for forensic applications in artificial intelligence.** *Proceedings of the American Academy of Forensic Sciences*, Orlando, FL, vol. 29, p. 380 (2023).
- Aggazzotti, Cristina. **A comparison of two computational approaches to analyzing last words.** *Proceedings of the American Academy of Forensic Sciences*, Seattle, WA, vol. 28, p. 483 (2022).
- Aggazzotti, Cristina. **A quantitative comparison of last words to related text genres.** *Proceedings of the American Academy of Forensic Sciences*, Seattle, WA, vol. 28, p. 486 (2022).
- Aggazzotti, Cristina. **A quantitative assessment of last words using Suicide Note Assessment REsearch (SNARE).** *Proceedings of the American Academy of Forensic Sciences*, held virtually, vol. 27, p. 370 (2021).
- Aggazzotti, Cristina and Stuart M. Shieber. **Reflexives and reciprocals in synchronous tree adjoining grammar.** *Proceedings of the 13th International Workshop on Tree Adjoining Grammars and Related Formalisms*, Association for Computational Linguistics, Umeå, Sweden, pp. 31–42 (2017).
- Aggazzotti, Cristina, Katherine Ramsland, Carole E. Chaski, Elizabeth Allyn Smith, Roderick T. Kennedy. **Communication Strategies To Minimize Bias and Strengthen Scientific Foundations in Forensic Science.** *Proceedings of the American Academy of Forensic Sciences*, New Orleans, LA, vol. 23, p. 38 (2017).
- Chaski, Carole E., Elizabeth Allyn Smith, and Cristina Aggazzotti. **Confirmation Bias and Metalinguistic Awareness.** *Proceedings of the American Academy of Forensic Sciences*, Las Vegas, NV, vol. 22, p. 902 (2016).

Skills

Computer	MS Excel (advanced); Python, GitHub, \LaTeX (intermediate); R, MATLAB, SPSS, OCaml (novice)
Linguistics	Syntax, semantics, corpus annotation and statistics, computational linguistics, tree adjoining grammar
Language	English (native); Italian, Spanish (intermediate); French (reading knowledge)
Soft skills	Adaptable, good communicator, detail-oriented, problem solver, team player, quick learner, enthusiastic

Interests

- Traveling (70 countries, 46 U.S. states), learning new skills and hobbies, adventuring, sailing, music, crafts, sports