

Khia A. Johnson

Psycholinguist, Speech Scientist & Ph.D. Candidate

Seattle, WA ·  · [@khia_johnson](https://www.linkedin.com/in/khiajohnson/) · <https://khiajohnson.com>

Summary

Linguistics Ph.D. candidate with 3+ years experience developing audio data sets, designing and conducting perception experiments, analyzing speech acoustics and survey data with quantitative methods using Python and R. Independently motivated, clear communicator, and strong technical writer with extensive presentation and publication record. Passionate about interdisciplinary collaboration, understanding human behavior, and using cognitive speech science to make social media a better place.

Experience

Project Lead, SpiCE: A corpus of Speech in Cantonese and English, Vancouver, BC 2018–present

- Created a new bilingual data set with high-quality audio recordings and time-aligned transcripts of semi-structured qualitative interviews; available via <https://spice-corpus.rtfid.io>
- Managed a team of 3 research assistants to develop interview materials, schedule and complete in-person sessions, create annotation guidelines and write documentation
- Designed an annotation pipeline to expedite manual transcription, leveraging speech-to-text, Python NLP libraries, and command-line tools, saving over 300 research assistant hours
- Measured, analyzed, and interpreted acoustic measurement data with corpus—using exploratory and statistical methods—to better understand multilingual voice quality and consonant production; includes work with voice onset time, consonant intensity, voicing, and sibilant spectral trajectories
- Analyzed corpus data in 4 projects, identifying samples, developing data quality checks, data cleaning, statistical modeling, causal inference, and technical writing

Research Assistant, Speech-in-Context Lab, Vancouver, BC

2017–present

- Programmed and analyzed lab and online speech perception experiments using yes/no, same/different, categorization, and goodness rating tasks in both within- and between-subjects designs with natural and synthesized speech stimuli
- Analyzed corpus and experimental data with R, using descriptive statistics, data visualization, principal components analyses, and mixed-effects models
- Designed manual annotation tasks with detailed guidelines for multiple projects, including acoustic landmark identification, token categorization, and goodness ratings
- Improved Qualtrics language background survey, for example, by implementing URL-based participant tracking for online experiments and revising demographic questions on sex and gender to be more appropriate and respectful

Selected Projects, University of British Columbia, Vancouver, BC

2019–present

- Designed, recorded, and processed acoustic stimuli for multi-day online phonetic learning experiment using an oddball detection paradigm
- Scraped social media text data and analyzed English language learners' use of intensifiers using Python, BeautifulSoup, NLTK, and multinomial logistic regression; available at <https://git.io/JWgAg>
- Developed a tutorial and led two workshops on creating online experiments using JavaScript, CSS, HTML, and hosting on AWS S3; available at <https://osf.io/3ufx7/>
- Maintains a blog with tutorials and tips for working with corpus language data using Python, R, SQLite, and audio analysis libraries; available at <https://khiajohnson.com/post>
- Presented research for a wide range of audiences, including at academic/industry conferences, a *PhDs Go Public* talk series, and for a Deep Learning and NLP reading group

Skills

Programming Python, R, SQLite, git, zsh/unix

Experiment design jsPsych (JavaScript library), PsychoPy

Transcription & analysis Praat, ELAN, VoiceSauce, Forced Alignment

Selected methods Bayesian and frequentist mixed effects models, Linear regression, Logistic regression, Growth curve modeling, Principal components analysis, Canonical correlation analysis, Inter-rater reliability, Descriptive statistics, t-test, Signal detection theory, Data simulation

Languages English (native), Spanish (fluent), Arabic (beginner)

Education

Ph.D. in Linguistics University of British Columbia, Vancouver, BC

2016–present

Linguistic Society of America Summer Institute University of California, Davis, CA

2019

B.A. in Linguistics Magna Cum Laude, University of Washington, Seattle, WA

2013