

## Yuan Chai

San Diego, CA · (303) 803-2798 · [yuc521@ucsd.edu](mailto:yuc521@ucsd.edu) · [github.com/ychaiucsd/source-code](https://github.com/ychaiucsd/source-code)

### EDUCATION

<b>University of California San Diego</b> <i>Ph.D. Linguistics</i> <b>Honors:</b> Teaching excellence award, Department of Linguistics	La Jolla, CA <b>GPA:</b> 3.98	June 2023
<b>University of Colorado, Boulder</b> <i>Master of Arts, Linguistics</i>	Boulder, CO <b>GPA:</b> 4.0	June 2017
<b>Beijing Normal University</b> <i>Bachelor of Arts, English</i> <b>Honors:</b> National fellowship; Honor graduate	Beijing, China <b>GPA:</b> 91.41/100	June 2015

### EXPERIENCE

<b>University of California San Diego</b> <i>Student researcher, Instructor, Teaching assistant</i>	La Jolla, CA	Sep, 2017 – Present
<ul style="list-style-type: none"><li>• <i>Phonetics of checked syllables in Xiapu Min:</i> Segment, annotate, and process 1 thousand sound files in Praat and VoiceSauce; clean and aggregate 200 thousand data points of speech sound signal in R; build linear mixed-effect models; conduct linear discriminant analysis; fit smoothed splines to time contours; create Praat scripts to manipulate fundamental frequency, duration, and amplitude of sound signal; design and execute speech production experiments in PsychoPy; build web-based speech perception experiments using HTML, CSS, and JavaScript; perform logistic regression and Bayesian analysis of forced-choice task results.</li><li>• <i>Field research on Rarámuri:</i> Annotate 6 thousand sound files in Praat and process them in VoiceSauce; clean and aggregate 20 thousand data points of speech signal output in R; visualize voice quality and fundamental frequency data using R::ggplot2</li><li>• <i>Voicing of glottal consonants and non-modal vowels:</i> Annotate, clean, and aggregate 230 thousand points of energy and voicing percentage data of sound signal from 201 languages in R</li><li>• <i>Comparing H1 vs. H1–H2 for representing voice quality:</i> Segmented and annotated 9 thousand sound files in R; clean and aggregate 990 thousand data points of speech signal output in R; analyze the correlation between electroglottography data and acoustic data in 9 languages; conduct model comparisons for parameter optimization</li><li>• <i>Popularity predictors for Airbnb listings in San Diego:</i> Clean data for 12 thousand Airbnb listings in San Diego; conduct lasso regressions between the occupancy rate of listings and 32 predictors; conduct parameter optimization by selecting 8 most effective predictors out of 32 predictors</li></ul>		
<b>University of Colorado, Boulder</b> <i>Student researcher, Teaching assistant, Research assistant</i>	Boulder, CO	Aug, 2015 - 2017
<ul style="list-style-type: none"><li>• <i>POS tagging:</i> Create Python scripts to randomly split corpus into training and test sets; build Hidden Markov Model in Python to assign part-of-speech tags to words; reach 93% accuracy in test set with 31 thousand words.</li><li>• <i>Sentiment analysis:</i> Use Naïve Bayes classifier and positive/negative lexicon from Bing Liu Opinion Lexicon as features to perform sentiment analysis on Movie Review Corpus by NLTK in Python; reach 82% accuracy in test</li><li>• <i>N-gram language models:</i> Split sentences into unigrams and bigrams in Python; calculate the probability of the sentences based on the frequency of the unigram and bigrams in the training corpus.</li><li>• <i>Interactive learning tool for relative clause:</i> Grep sentences with relative clauses from Brown Corpus in NLTK, extract the relative pronouns out of the sentences; build a fill-in-blank interface in Python for users to fill in the relative pronoun of relative clauses and get feedback.</li></ul>		

### SKILLS

<b>Data manipulation &amp; analytics:</b>	R::tidyverse, MASS, lme4, pspline, brms, rjags; Python; SQL
<b>Visualization:</b>	R::ggplot, phonR
<b>Web development:</b>	HTML, CSS, JavaScript
<b>Speech processing</b>	Praat, VoiceSauce, ELAN, PsychoPy

### PUBLICATIONS

- **Chai, Y.** 2019. The Source of Creak in Mandarin. *Proceedings of ICPhS 2019*.
- **Chai, Y.**, Kunda, T., Rodríguez, A. & Rose, S. 2022. The prosody of declaratives and questions in Rere. In H. Kubozono, J. Ito & A. Mester (Eds.), *Prosody and Prosodic Interfaces*, Oxford: Oxford University Press.
- Caballero, G., **Chai, Y.** & Garellek, M. 2022. Stress, tone, and intonation in Choguita Rarámuri. In H. Kubozono, J. Ito & A. Mester (Eds.), *Prosody and Prosodic Interfaces*, Oxford: Oxford University Press.