|  |
| --- |
| Miquel Simonet1, Joseph V. Casillas, & Alex Aldrich3 |
| 1 University of Arizona |
| 2 Rutgers University |
| 3 Concordia College |
|  |

# Author note

Correspondence concerning this article should be addressed to Miquel Simonet, Postal address. E-mail: [my@email.com](mailto:my@email.com)

# Overview

## Bayesian data analysis

This study employs Bayesian Data Analysis for quantitative inferential statistics. Specifically, this implies that we use Bayesian *credible intervals* to draw statistical inferences. A Bayesian model calculates a posterior distribution, i.e., a distribution of plausible parameter values, given the data, a data-generating model, and any prior information we have about those parameter values. Posterior distributions are computationally costly. For this reason, we use the Hamiltonian Markov Chain Monte Carlo algorithm to obtain a sample that incldues thousands of values from the posterior distribution. In practical terms, what this means is that we do not calculate a single point estimate for an effect , but rather we draw a sample of 4,000 plausible values for . This allows us to quantify our uncertainty regarding by summaryzing the distribution of those values. We will use 3 statistics to describe the posterior distribution: (1) the mean, (2) a credible interval, and (3) the posterior probability that the effect is greater than 0. The mean provides a point estimate for the distribution, the 95% credible interval provides bounds for the effect, and the posterior probability that the effect is greater than zero further quantifies our uncertainty. For example, if a hypothesis states that , we judge there to be *compelling evidence* for this hypothesis if the mean point estimate is a positive number, is (by a reasonably clear margin) not included in the 95% credible interval of , and the posterior is close to one.

These three statistics provide an intuitive interpretation of any given effect. For instance, consider a case in which the posterior mean of is 100 and the 95% credible interval is [40, 160]. The interval tells us that we can be 95% certain the *true* value of is between 40 and 160, given the data, our model, and our prior information. Furthermore, the interval allows us to specify areas of uncertainty. In this example, we can conclude that the effect is almost certain to be positive. The lower interval value of 40 tells us that 95% of the plausible values are greater than 40. We also note that the interval covers a wide range of values, thus we also conclude that we are not very certain about the size of the effect. This type of interepretation is not possible under a frequentist paradigm.

## About the data

### Metrics

### Participants

The data include 3 populations: monolingual English speakers, monolingual Spanish speakers, Bilingual English-Spanish speakers.

#### Monolingual English speakers

The monolingual English speakers are from the U.S. Southwest. Age and other info here.

#### Monolingual Spanish speakers

The monolingual Spanish speakers com from Majorca, Spain. They are actual bilingual. There are N of them. Age and other info here.

#### Bilingual speakers

The English-Spanish bilinguals are from Southern Arizona, Northern Mexico. Age, BLP and other info. Note there are two samples from this population.

## About this document

This document was written in RMarkdown using papaja and serves as a project report for our research group. From this point foward, the document is written as if it were the results section of a future manuscript. This implies that it is written in a way that allows it to be copy and pasted into the actual manuscript once it is available.

# Results

## Experiment 1: Monolinguals

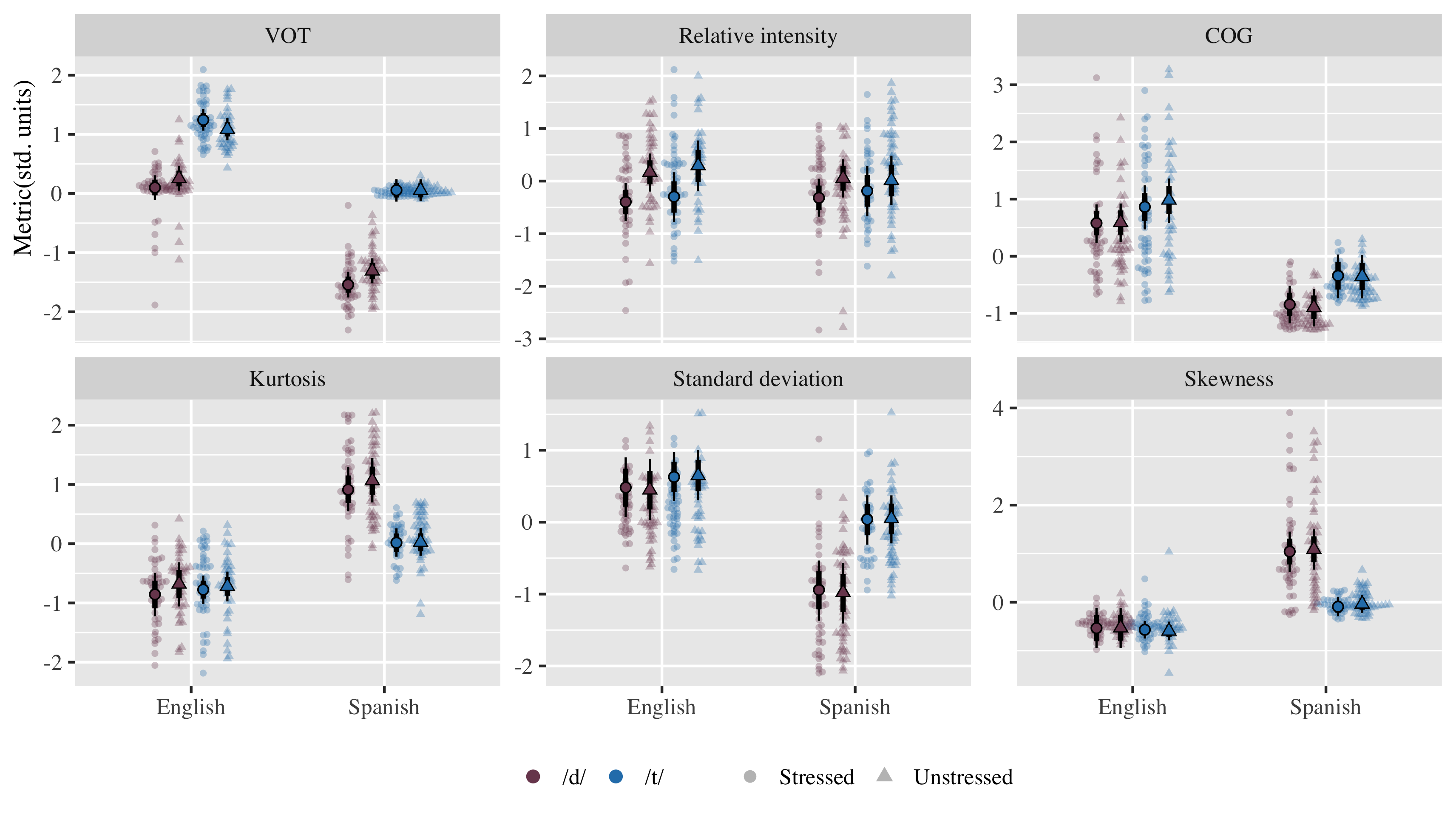


Figure 1: VOT and burst metrics of coronal stops (/d/, /t/) from monolingual speakers as a function of language (English, Spanish), and lexical stress (Stressed, Unstressed). Transparent points represent raw data. Solid points indicate posterior means ± 95% and 80% credible intervals.

## Experiment 2: Bilinguals

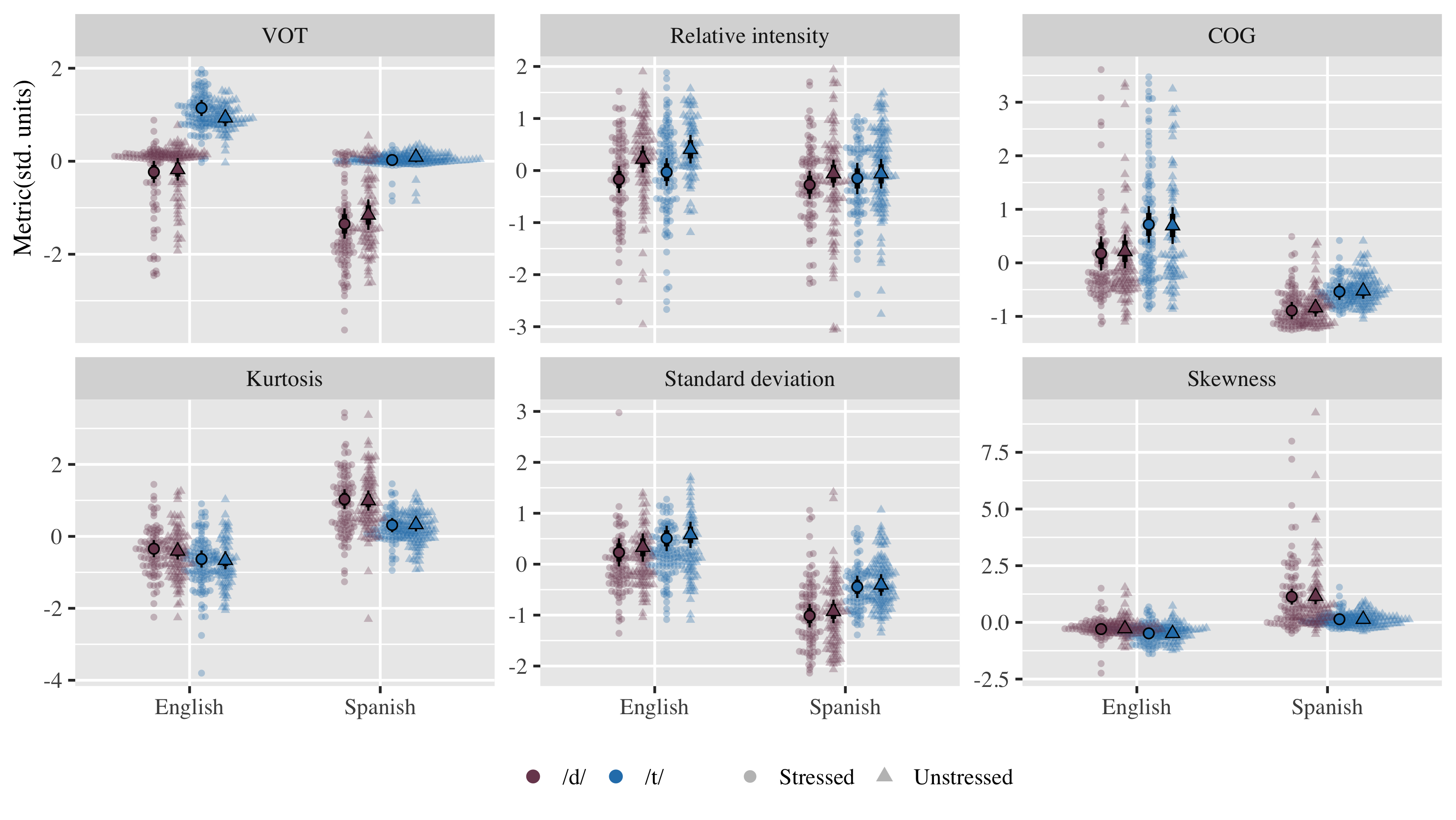


Figure 2: VOT and burst metrics of coronal stops (/d/, /t/) from bilingual speakers as a function of language (English, Spanish), and lexical stress (Stressed, Unstressed). Transparent points represent raw data. Solid points indicate posterior means ± 95% and 80% credible intervals.

## Experiment 3: Bilingual POA data

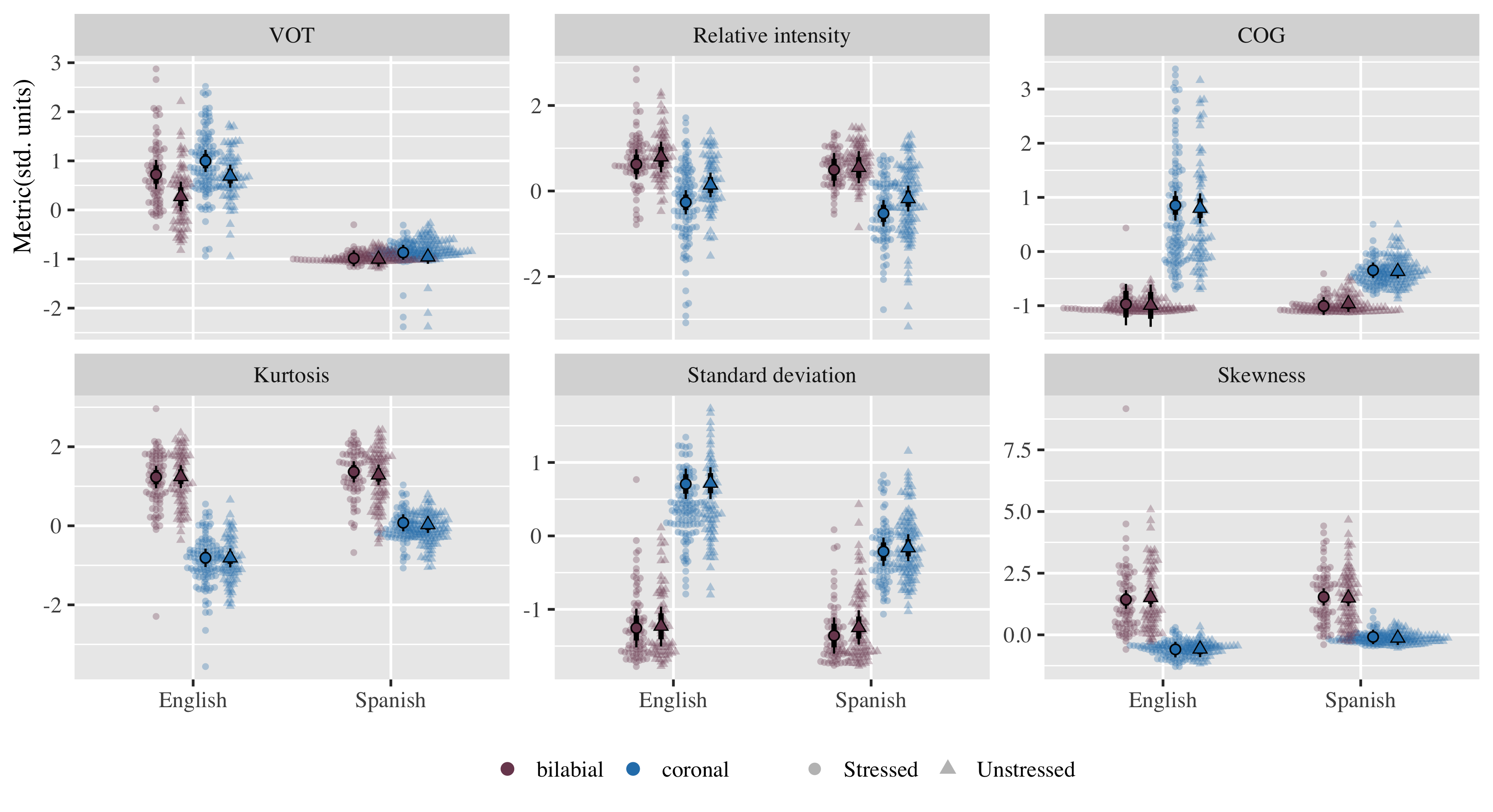


Figure 3: VOT and burst metrics of voiceless stops from bilingual speakers as a function of language (English, Spanish), place of articulation (Coronal, Bilabial), and lexical stress (Stressed, Unstressed). Transparent points represent raw data. Solid points indicate posterior means ± 95% and 80% credible intervals.