|  |
| --- |
| 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, 1423 E. University Blvd. Rm. 594 Modern Languages Building, Tucson, Arizona 85721. E-mail: [simonet@email.arizona.edu](mailto:simonet@email.arizona.edu)

Abstract

Here.

*Keywords:* Spanish, Coronal stops, Spectral moments

*Word count:*

# Method

## Participants

The data include 42 participants from 3 populations: monolingual English speakers, monolingual Spanish speakers, and bilingual Spanish-English speakers. All participants were females between the ages of 18 and 23. The monolingual English speakers were recorded in English and the monolingual Spanish speakers were recorded in Spanish. The Spanish-English bilinguals were recorded in both of their languages.

**Monolingual English speakers.** The study includes 8 monolingual English speakers. They were undergraduate students at the University of Arizona, born and raised in the US Southwest. The English speakers were functionally monolingual, though they reported having taken introductory Spanish courses. They were not able to maintain a basic conversation in Spanish. All of the participants in this group reported English as their native language and verified not having been exposed to any other languages while growing up.

**Monolingual Spanish speakers.** The monolingual Spanish group comprised 8 speakers that were recruited from the *Universitat de les Illes Balears* campus community and were born and raised on the island of Majorca, Spain. They reported that, although they had studied some English in Spain, they were not able to maintain a basic conversation in this language. The participants of this group also speak Catalan. Importantly, there are no reported differences in the phonetic realization of voice timing between the Spanish and Catalan, nor are there place difference between the coronal stops.

**Bilingual speakers.** The English-Spanish bilinguals (n = 26) came from Southern Arizona and Northern Mexico. There are two samples from this population. The coronal dataset includes 17 speakers and the bilabial dataset includes 9 speakers. The Spanish-English bilinguals were undergraduate students at the University of Arizona in Tucson, Arizona. The bilinguals were brought up by Spanish-speaking families and were schooled mostly in English. They reported using English and Spanish daily, both in the classroom as well as with their friends and relatives.

## Metrics

F1/F2, voice onset time, relative intensity, center of gravity, standard deviation, skewness, kurtosis

## Procedure

Decide if we will present 3 separate experiments with 3 different methods sections.

## Statistical analyses

All analyses were conducted in R (R Core Team, 2019, version 3.6.0). We use Bayesian multilevel models fitted in Stan using brms (Bürkner, 2017, 2018, version 2.10.0). Bayesian Data Analysis (BDA) is an alternative to frequentist statistical analysis.[[1]](#footnote-27) For all models, the criterion was standardized, or converted to z-scores, in order to facilitate comparability between metrics. Continuous predictors were also standardized and categorical predictors were sum-to-zero coded. Thus for all models the intercept represents the outcome variable at the grand mean in standardized units. We used regularizing, weakly informative priors in all models (specifics below) with 4,000 iterations (2,000 warm-up) running on 16 processing cores. We quantify our uncertainty regarding a given effect by reporting point estimates derived from the posterior predictive distribution, including the 95% Highest Density Credible Intervals (HDI). Additionally, we assume a negligible effect size of ± 0.1 (Cohen, 1988, 2013; Kruschke, 2018) in order to establish a Region of Practical Equivalence (ROPE), for which we assess the proportion of the HDI that falls within this interval. Finally, we report the Maximum Probability of Effect (MPE), or the Probability of Direction, as the proportion of the posterior distribution that is of the median’s sign. We assume there to be compelling evidence for a given effect when the HDI of the posterior distribution does not contain 0 nor fall within the ROPE by a reasonable margin and the MPE is close to 1.

# Results

The results are divided into 3 subsections dealing with (1) monolingual data, (2) bilingual data, and (3) POA data. Each subsection presents the results of 6 metrics: VOT, relative intensity, center of gravity, kurtosis, standard deviation, and skewness. We report only relevant effects. Please see the supplementary materials (Appendices A-D) for complete model summaries.

## Experiment 1: Monolinguals

We modeled VOT and the burst metrics as a function of language (English, Spanish), phoneme (/d/, /t/), standardized F1 and F2, and item repetitions. The model used regularizing, weakly informative priors . Specifically, all parameters were assumed to be distributed as normal with a standard deviation of 5, i.e., , . The standard deviation parameters for random effects and the model residual error (sigma) were truncated to exclude negative values. Figure 1 plots VOT and the burst metrics as a function of language (English, Spanish) and phoneme (/d/, /t/). For all plots the y-axis represents the outcome variable in standardized units. The x-axis indicates the language and voiced/voiceless pairs are represented by color.

![Figure 1: VOT and burst metrics of coronal stops (/d/, /t/) from monolingual speakers as a function of language (English, Spanish). Transparent points represent raw data. Solid points indicate posterior means ± 99% and 80% credible intervals.](data:application/pdf;base64,)

Figure 1: VOT and burst metrics of coronal stops (/d/, /t/) from monolingual speakers as a function of language (English, Spanish). Transparent points represent raw data. Solid points indicate posterior means ± 99% and 80% credible intervals.

**Voice-onset time**. English stops had higher VOT than Spanish stops (β = 0.67; SE = 0.05; CI = [0.57, 0.77]) and the voiceless segments had higher VOT than the voiced segments (β = −0.62; SE = 0.04; CI = [−0.69, −0.54]). The HDI’s for language and phoneme were completely outside the ROPE and the MPE was 100%. Thus there was compelling evidence for an effect of both language and phoneme, as well as a moderate interaction between the two predictors (β = 0.11; SE = 0.03; CI = [0.04, 0.18]), such that the voicing difference between Spanish coronals was slightly larger than that of the English coronals. For the language × phoneme interaction 36.38% of the HDI fell within the ROPE and the MPE was .998. Additionally, we compared the short-lag stops of each language, English /d/ and Spanish /t/, and found no envidence that the segments differed from each other (β = 0.1; CI = [-0.152, 0.345]; ROPE = 0.479; MPE = 0.795). The VOT data is plotted in the first panel of Figure 1. The complete model summary and the short-lag stop comparison are available in Table 2 and Figure 9 of the supplementary materials.

**Relative intensity**. The relative intensity data is plotted in the top middle panel of Figure 1. English and Spanish stops differed very little with regard to relative intensity. The model provided no compelling evidence for a language effect (β = 0.08; SE = 0.11; CI = [−0.14, 0.29]), nor for phoneme (β = −0.10; SE = 0.07; CI = [−0.23, 0.04]) or vowel height effects (β = −0.11; SE = 0.06; CI = [−0.22, −0.00]). Frontness of the following vowel did modulate relative intensity of the burst (β = −0.22; SE = 0.06; CI = [−0.34, −0.11]) such that higher F2 values were associated with lower relative intensity. Finally, there was no compelling evidence for a language × phoneme interaction (β = 0.02; SE = 0.07; CI = [−0.11, 0.15]).

**Center of gravity.** The center of gravity (COG) data is plotted in the third panel of Figure 1. English stops had a higher COG than Spanish stops (β = 0.64; SE = 0.12; CI = [0.40, 0.87]), and the voiceless segments had a higher COG than the voiceled segments (β = −0.21; SE = 0.03; CI = [−0.28, −0.14]). In both cases the HDI’s were completely outside the ROPE and the MPE was 100%. This suggests compelling evidence for an effect of both language and phoneme. The two predictors did not interact (β = 0.06; SE = 0.03; CI = [−0.01, 0.13]), nor did F1 (β = −0.02; SE = 0.03; CI = [−0.07, 0.03]) or F2 (β = 0.04; SE = 0.03; CI = [−0.03, 0.09]) have any influence on COG.

**Kurtosis.**

(β = −0.62; SE = 0.08; CI = [−0.77, −0.46])

(β = 0.22; SE = 0.05; CI = [0.11, 0.32])

(β = −0.03; SE = 0.03; CI = [−0.09, 0.03])

(β = −0.01; SE = 0.03; CI = [−0.07, 0.05])

(β = −0.24; SE = 0.05; CI = [−0.35, −0.13])

**Standard deviation.**

(β = 0.50; SE = 0.11; CI = [0.28, 0.72])

(β = −0.28; SE = 0.06; CI = [−0.40, −0.16])

(β = −0.00; SE = 0.03; CI = [−0.06, 0.05])

(β = 0.02; SE = 0.03; CI = [−0.05, 0.08])

(β = 0.20; SE = 0.06; CI = [0.09, 0.32])

**Skewness.**

(β = −0.50; SE = 0.07; CI = [−0.64, −0.37])

(β = 0.30; SE = 0.07; CI = [0.16, 0.43])

(β = −0.04; SE = 0.03; CI = [−0.11, 0.02])

(β = 0.00; SE = 0.03; CI = [−0.05, 0.06])

(β = −0.25; SE = 0.07; CI = [−0.38, −0.12])

![Figure 2: Posterior medians ± 95% and 66% credible intervals for VOT and burst metrics of monolingual coronal stops.](data:application/pdf;base64,)

Figure 2: Posterior medians ± 95% and 66% credible intervals for VOT and burst metrics of monolingual coronal stops.

## Experiment 2: Bilinguals

Model info here.

![Figure 3: VOT and burst metrics of coronal stops (/d/, /t/) from bilingual speakers as a function of language (English, Spanish). Transparent points represent raw data. Solid points indicate posterior means ± 99% and 80% credible intervals.](data:application/pdf;base64,)

Figure 3: VOT and burst metrics of coronal stops (/d/, /t/) from bilingual speakers as a function of language (English, Spanish). Transparent points represent raw data. Solid points indicate posterior means ± 99% and 80% credible intervals.

**Voice-onset time.**

(β = 0.48; SE = 0.04; CI = [0.39, 0.56])

(β = −0.62; SE = 0.06; CI = [−0.74, −0.50])

(β = 0.00; SE = 0.02; CI = [−0.03, 0.03])

(β = −0.01; SE = 0.02; CI = [−0.04, 0.02])

(β = −0.01; SE = 0.04; CI = [−0.09, 0.07])

**Relative intensity.**

(β = 0.14; SE = 0.05; CI = [0.03, 0.24])

(β = −0.08; SE = 0.04; CI = [−0.16, −0.01])

(β = −0.27; SE = 0.05; CI = [−0.38, −0.17])

(β = −0.19; SE = 0.03; CI = [−0.25, −0.14])

(β = 0.00; SE = 0.03; CI = [−0.06, 0.07])

**Center of gravity.**

(β = 0.58; SE = 0.07; CI = [0.44, 0.72])

(β = −0.22; SE = 0.03; CI = [−0.27, −0.16])

(β = 0.02; SE = 0.02; CI = [−0.02, 0.06])

(β = −0.02; SE = 0.02; CI = [−0.06, 0.02])

(β = −0.05; SE = 0.03; CI = [−0.11, 0.02])

**Kurtosis.**

(β = −0.60; SE = 0.05; CI = [−0.69, −0.50])

(β = 0.25; SE = 0.04; CI = [0.18, 0.32])

(β = −0.00; SE = 0.03; CI = [−0.05, 0.05])

(β = 0.01; SE = 0.03; CI = [−0.04, 0.06])

(β = −0.11; SE = 0.05; CI = [−0.21, −0.01])

**Standard deviation.**

(β = 0.56; SE = 0.06; CI = [0.43, 0.68])

(β = −0.22; SE = 0.04; CI = [−0.29, −0.14])

(β = 0.00; SE = 0.02; CI = [−0.04, 0.05])

(β = −0.01; SE = 0.02; CI = [−0.06, 0.03])

(β = 0.07; SE = 0.04; CI = [−0.01, 0.16])

**Skewness.**

(β = −0.51; SE = 0.04; CI = [−0.60, −0.43])

(β = 0.31; SE = 0.04; CI = [0.22, 0.39])

(β = −0.01; SE = 0.02; CI = [−0.05, 0.04])

(β = −0.00; SE = 0.02; CI = [−0.05, 0.04])

(β = −0.20; SE = 0.05; CI = [−0.30, −0.11])

![Figure 4: Posterior medians ± 95% and 66% credible intervals for VOT and burst metrics of bilingual coronal stops.](data:application/pdf;base64,)

Figure 4: Posterior medians ± 95% and 66% credible intervals for VOT and burst metrics of bilingual coronal stops.

## Experiment 3: Bilingual POA data

Model info here.

![Figure 5: VOT and burst metrics of voiceless stops from bilingual speakers as a function of language (English, Spanish), place of articulation (Coronal, Bilabial). Transparent points represent raw data. Solid points indicate posterior means ± 99% and 80% credible intervals.](data:application/pdf;base64,)

Figure 5: VOT and burst metrics of voiceless stops from bilingual speakers as a function of language (English, Spanish), place of articulation (Coronal, Bilabial). Transparent points represent raw data. Solid points indicate posterior means ± 99% and 80% credible intervals.

**Voice-onset time.**

(β = 0.77; SE = 0.05; CI = [0.67, 0.88])

(β = 0.14; SE = 0.05; CI = [0.04, 0.25])

(β = 0.02; SE = 0.02; CI = [−0.01, 0.05])

(β = −0.01; SE = 0.02; CI = [−0.04, 0.03])

(β = 0.03; SE = 0.05; CI = [−0.08, 0.13])

**Relative intensity.**

(β = 0.12; SE = 0.04; CI = [0.04, 0.19])

(β = −0.41; SE = 0.11; CI = [−0.63, −0.19])

(β = −0.22; SE = 0.05; CI = [−0.31, −0.13])

(β = −0.18; SE = 0.04; CI = [−0.25, −0.10])

(β = 0.03; SE = 0.04; CI = [−0.05, 0.11])

**Center of gravity.**

(β = 0.30; SE = 0.05; CI = [0.20, 0.41])

(β = 0.65; SE = 0.06; CI = [0.52, 0.77])

(β = 0.04; SE = 0.02; CI = [−0.01, 0.09])

(β = −0.04; SE = 0.02; CI = [−0.08, 0.01])

(β = 0.29; SE = 0.06; CI = [0.18, 0.40])

**Kurtosis.**

(β = −0.23; SE = 0.04; CI = [−0.30, −0.16])

(β = −0.83; SE = 0.07; CI = [−0.97, −0.68])

(β = 0.00; SE = 0.02; CI = [−0.04, 0.05])

(β = 0.01; SE = 0.02; CI = [−0.04, 0.06])

(β = −0.19; SE = 0.03; CI = [−0.26, −0.12])

**Standard deviation.**

(β = 0.23; SE = 0.05; CI = [0.14, 0.32])

(β = 0.78; SE = 0.06; CI = [0.66, 0.90])

(β = 0.01; SE = 0.02; CI = [−0.04, 0.06])

(β = −0.02; SE = 0.02; CI = [−0.07, 0.03])

(β = 0.20; SE = 0.04; CI = [0.12, 0.29])

**Skewness.**

(β = −0.13; SE = 0.05; CI = [−0.23, −0.03])

(β = −0.87; SE = 0.09; CI = [−1.04, −0.68])

(β = 0.01; SE = 0.03; CI = [−0.06, 0.08])

(β = 0.01; SE = 0.02; CI = [−0.03, 0.06])

(β = −0.09; SE = 0.05; CI = [−0.19, 0.01])

![Figure 6: plot-poa-bilinguals-summary](data:application/pdf;base64,)

Figure 6: plot-poa-bilinguals-summary

# Discussion

## Summary of findings

## Interpretation and implications

# References

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Schoot, R. van de, & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. *European Health Psychologist*, *16*(2), 75–84.

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, *71*, 147–161.

Appendix A

## Experiment 0: Vowels

Prior to the stop analyses we scrutinezed the formant structure, F1 and F2, of the vowel following the coronal stops. The purpose of this analysis was to determine if the low /a/ and /ae/ vowels of Spanish and English, respectively, differed from each other. This analysis was taken as a precautionary measure with the objective of determining whether or not coarticulation effects may be present in the stop metrics due to the possible different spectral envelopes of the next segment.

The F1 and F2 data were analayzed using separate Bayesian multilevel models. The criterion (F1 or F2) were standardized and modeled as a function of language (English, Spanish), phoneme (/d/, /t/), and item repetition. Language and phoneme were sum coded (-1, 1). The random effects structure included by-subject intercepts with random slopes for phoneme and item repetition, as well as by-item intercepts with a random slope for item repetition. The model included weakly informative priors with the mean centered at 0 and a standard deviation of 2.

The model suggested weak evidence for a language effect on F1 (β = 0.18; SE = 0.09; CI = [0.01, 0.36]) and F2 (β = 0.20; SE = 0.11; CI = [−0.03, 0.42]). For F1 and F2, 15.41% and 15.41% of the respective HDI’s fell within the specified ROPE. The probability that β was greater than 0 was .979 for F1 and .958 for F2. Together, the point estimates suggest the spectral centroid of the Spanish vowel was slightly higher and more posterior with respect to that of the English vowel. There was no evidence for a phoneme effect on F1 (β = −0.02; SE = 0.04; CI = [−0.10, 0.05]), nor on F2 (β = −0.01; SE = 0.04; CI = [−0.09, 0.08]). In both cases 100% of the HDI fell within the ROPE and the probability that β was less than 0 was .717 for F1 and .555 for F2. Given the possibility of a small effect of language on the spectal envelope, standardized F1 and F2 were used in subsequent analyses of the coronal stops to control for any coarticulatory effects on the burst. Figure 7 plots the F1 × F2 data and Figure 8 plots the model summary. A complete summary of the F1 and F2 models is available in Table 1.

![Figure 7: F1 × F2 of /a/ from monolingual speakers as a function of language (English, Spanish). Transparent points represent raw data. Solid points indicate posterior means ± 95% and 80% credible intervals.](data:application/pdf;base64,)

Figure 7: F1 × F2 of /a/ from monolingual speakers as a function of language (English, Spanish). Transparent points represent raw data. Solid points indicate posterior means ± 95% and 80% credible intervals.

![Figure 8: Posterior medians ± 95% and 66% credible intervals for F1 and F2 from monolingual speaker data.](data:application/pdf;base64,)

Figure 8: Posterior medians ± 95% and 66% credible intervals for F1 and F2 from monolingual speaker data.

Table 1: Model summary for F1 and F2 as a function of language (English, Spanish), phoneme (/d/, /t/), and item repetition for monolingual data. The percentage of the HDI contained within the ROPE is based on an effect size of ±0.1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Metric | Parameter | Estimate | HDI | ROPE | MPE |
| F1 | Intercept | −0.014 | [−0.245, 0.219] | 0.650 | 0.550 |
|  | Language | 0.184 | [−0.001, 0.353] | 0.154 | 0.979 |
|  | Phoneme | −0.021 | [−0.096, 0.053] | 1.000 | 0.717 |
|  | Item rep. | 0.008 | [−0.074, 0.091] | 1.000 | 0.573 |
| F2 | Intercept | −0.094 | [−0.368, 0.176] | 0.467 | 0.759 |
|  | Language | 0.196 | [−0.037, 0.409] | 0.181 | 0.958 |
|  | Phoneme | −0.006 | [−0.090, 0.082] | 1.000 | 0.555 |
|  | Item rep. | 0.045 | [−0.033, 0.123] | 0.938 | 0.874 |

Appendix B

Table 2: Model summary for VOT and burst metrics as a function of language (English, Spanish), phoneme (/d/, /t/), F1, F2, and item repetition for monolingual coronal stops. The percentage of the HDI contained within the ROPE is based on an effect size of ± 0.1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Metric | Parameter | Estimate | HDI | ROPE | MPE |
| VOT | Intercept | 0.011 | [−0.124, 0.129] | 0.926 | 0.575 |
|  | Language | 0.666 | [0.567, 0.767] | 0.000 | 1.000 |
|  | Phoneme | −0.616 | [−0.691, -0.548] | 0.000 | 1.000 |
|  | F1 | 0.01 | [−0.030, 0.052] | 1.000 | 0.704 |
|  | F2 | −0.006 | [−0.048, 0.035] | 1.000 | 0.626 |
|  | Item rep. | −0.013 | [−0.044, 0.019] | 1.000 | 0.790 |
|  | Language × Phoneme | 0.111 | [0.042, 0.178] | 0.364 | 0.998 |
| RI | Intercept | −0.12 | [−0.392, 0.146] | 0.409 | 0.821 |
|  | Language | 0.077 | [−0.142, 0.290] | 0.564 | 0.768 |
|  | Phoneme | −0.097 | [−0.227, 0.038] | 0.518 | 0.925 |
|  | F1 | −0.113 | [−0.222, -0.001] | 0.393 | 0.977 |
|  | F2 | −0.221 | [−0.334, -0.105] | 0.000 | 1.000 |
|  | Item rep. | 0.058 | [−0.025, 0.147] | 0.855 | 0.914 |
|  | Language × Phoneme | 0.022 | [−0.108, 0.148] | 0.904 | 0.635 |
| COG | Intercept | 0.048 | [−0.195, 0.296] | 0.593 | 0.658 |
|  | Language | 0.638 | [0.410, 0.880] | 0.000 | 1.000 |
|  | Phoneme | −0.209 | [−0.276, -0.141] | 0.000 | 1.000 |
|  | F1 | −0.015 | [−0.065, 0.035] | 1.000 | 0.728 |
|  | F2 | 0.035 | [−0.025, 0.095] | 1.000 | 0.881 |
|  | Item rep. | −0.028 | [−0.074, 0.019] | 1.000 | 0.878 |
|  | Language × Phoneme | 0.058 | [−0.007, 0.127] | 0.916 | 0.957 |
| Kurtosis | Intercept | −0.117 | [−0.297, 0.072] | 0.430 | 0.901 |
|  | Language | −0.615 | [−0.781, -0.469] | 0.000 | 1.000 |
|  | Phoneme | 0.216 | [0.113, 0.327] | 0.000 | 1.000 |
|  | F1 | −0.028 | [−0.089, 0.034] | 1.000 | 0.823 |
|  | F2 | −0.008 | [−0.063, 0.051] | 1.000 | 0.608 |
|  | Item rep. | 0.013 | [−0.050, 0.073] | 1.000 | 0.668 |
|  | Language × Phoneme | −0.24 | [−0.348, -0.136] | 0.000 | 1.000 |
| SD | Intercept | 0.036 | [−0.213, 0.278] | 0.603 | 0.627 |
|  | Language | 0.497 | [0.281, 0.710] | 0.000 | 1.000 |
|  | Phoneme | −0.282 | [−0.397, -0.162] | 0.000 | 1.000 |
|  | F1 | −0.005 | [−0.063, 0.054] | 1.000 | 0.560 |
|  | F2 | 0.018 | [−0.041, 0.086] | 1.000 | 0.719 |
|  | Item rep. | −0.027 | [−0.093, 0.040] | 1.000 | 0.798 |
|  | Language × Phoneme | 0.205 | [0.087, 0.315] | 0.016 | 0.999 |
| Skewness | Intercept | −0.025 | [−0.197, 0.155] | 0.768 | 0.608 |
|  | Language | −0.502 | [−0.642, -0.370] | 0.000 | 1.000 |
|  | Phoneme | 0.296 | [0.167, 0.431] | 0.000 | 1.000 |
|  | F1 | −0.044 | [−0.108, 0.024] | 0.979 | 0.911 |
|  | F2 | 0.005 | [−0.050, 0.061] | 1.000 | 0.569 |
|  | Item rep. | 0.017 | [−0.036, 0.069] | 1.000 | 0.731 |
|  | Language × Phoneme | −0.251 | [−0.377, -0.123] | 0.000 | 1.000 |

![Figure 9: Posterior distribution comparing the short-lag stops, English /d/ and Spanish /t/. The white point represents the posterior mean ± 95% HDI and the grey region represents the ROPE (± 0.1).](data:application/pdf;base64,)

Figure 9: Posterior distribution comparing the short-lag stops, English /d/ and Spanish /t/. The white point represents the posterior mean ± 95% HDI and the grey region represents the ROPE (± 0.1).

Appendix C

Table 3: Model summary for VOT and burst metrics as a function of language (English, Spanish), phoneme (/d/, /t/), F1, F2, and item repetition for bilingual coronal stops. The percentage of the HDI contained within the ROPE is based on an effect size of ± 0.1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Metric | Parameter | Estimate | HDI | ROPE | MPE |
| VOT | Intercept | −0.075 | [−0.216, 0.072] | 0.638 | 0.850 |
|  | Language | 0.476 | [0.391, 0.565] | 0.000 | 1.000 |
|  | Phoneme | −0.622 | [−0.737, -0.497] | 0.000 | 1.000 |
|  | F1 | 0.001 | [−0.028, 0.031] | 1.000 | 0.536 |
|  | F2 | −0.009 | [−0.041, 0.021] | 1.000 | 0.705 |
|  | Item rep. | −0.003 | [−0.034, 0.026] | 1.000 | 0.575 |
|  | Language × Phoneme | −0.007 | [−0.085, 0.070] | 1.000 | 0.565 |
| RI | Intercept | −0.025 | [−0.226, 0.184] | 0.706 | 0.598 |
|  | Language | 0.135 | [0.026, 0.237] | 0.240 | 0.994 |
|  | Phoneme | −0.084 | [−0.158, -0.014] | 0.677 | 0.988 |
|  | F1 | −0.271 | [−0.377, -0.166] | 0.000 | 1.000 |
|  | F2 | −0.193 | [−0.248, -0.138] | 0.000 | 1.000 |
|  | Item rep. | 0.004 | [−0.043, 0.052] | 1.000 | 0.566 |
|  | Language × Phoneme | 0.005 | [−0.062, 0.071] | 1.000 | 0.555 |
| COG | Intercept | −0.107 | [−0.286, 0.073] | 0.464 | 0.885 |
|  | Language | 0.581 | [0.441, 0.721] | 0.000 | 1.000 |
|  | Phoneme | −0.218 | [−0.272, -0.164] | 0.000 | 1.000 |
|  | F1 | 0.02 | [−0.023, 0.063] | 1.000 | 0.821 |
|  | F2 | −0.021 | [−0.064, 0.018] | 1.000 | 0.842 |
|  | Item rep. | 0.02 | [−0.019, 0.055] | 1.000 | 0.852 |
|  | Language × Phoneme | −0.047 | [−0.110, 0.016] | 0.977 | 0.929 |
| Kurtosis | Intercept | 0.066 | [−0.095, 0.226] | 0.675 | 0.797 |
|  | Language | −0.596 | [−0.693, -0.498] | 0.000 | 1.000 |
|  | Phoneme | 0.254 | [0.182, 0.322] | 0.000 | 1.000 |
|  | F1 | 0 | [−0.051, 0.052] | 1.000 | 0.503 |
|  | F2 | 0.005 | [−0.044, 0.055] | 1.000 | 0.584 |
|  | Item rep. | −0.044 | [−0.089, 0.003] | 1.000 | 0.969 |
|  | Language × Phoneme | −0.11 | [−0.209, -0.012] | 0.408 | 0.984 |
| SD | Intercept | −0.127 | [−0.285, 0.025] | 0.363 | 0.948 |
|  | Language | 0.556 | [0.436, 0.680] | 0.000 | 1.000 |
|  | Phoneme | −0.22 | [−0.296, -0.147] | 0.000 | 1.000 |
|  | F1 | 0.003 | [−0.045, 0.051] | 1.000 | 0.548 |
|  | F2 | −0.013 | [−0.061, 0.031] | 1.000 | 0.706 |
|  | Item rep. | 0.028 | [−0.015, 0.070] | 1.000 | 0.909 |
|  | Language × Phoneme | 0.074 | [−0.011, 0.165] | 0.740 | 0.953 |
| Skewness | Intercept | 0.122 | [−0.041, 0.276] | 0.377 | 0.935 |
|  | Language | −0.514 | [−0.597, -0.427] | 0.000 | 1.000 |
|  | Phoneme | 0.306 | [0.225, 0.394] | 0.000 | 1.000 |
|  | F1 | −0.01 | [−0.056, 0.035] | 1.000 | 0.663 |
|  | F2 | −0.003 | [−0.047, 0.044] | 1.000 | 0.543 |
|  | Item rep. | −0.009 | [−0.057, 0.039] | 1.000 | 0.651 |
|  | Language × Phoneme | −0.202 | [−0.297, -0.113] | 0.000 | 1.000 |

Appendix D

Table 4: Model summary for VOT and burst metrics as a function of language (English, Spanish), place of articulation (bilabial, coronal), F1, F2, and item repetition for bilingual voiceless stops. The percentage of the HDI contained within the ROPE is based on an effect size of ± 0.1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Metric | Parameter | Estimate | HDI | ROPE | MPE |
| VOT | Intercept | −0.094 | [−0.212, 0.023] | 0.550 | 0.944 |
|  | Language | 0.773 | [0.663, 0.878] | 0.000 | 1.000 |
|  | Place | 0.14 | [0.032, 0.242] | 0.211 | 0.996 |
|  | F1 | 0.017 | [−0.014, 0.048] | 1.000 | 0.853 |
|  | F2 | −0.006 | [−0.040, 0.029] | 1.000 | 0.628 |
|  | Item rep. | 0.002 | [−0.037, 0.039] | 1.000 | 0.538 |
|  | Language × Place | 0.025 | [−0.079, 0.133] | 0.934 | 0.686 |
| RI | Intercept | 0.211 | [−0.009, 0.428] | 0.139 | 0.970 |
|  | Language | 0.115 | [0.033, 0.192] | 0.341 | 0.997 |
|  | Place | −0.41 | [−0.622, -0.179] | 0.000 | 1.000 |
|  | F1 | −0.22 | [−0.309, -0.126] | 0.000 | 1.000 |
|  | F2 | −0.176 | [−0.252, -0.103] | 0.000 | 1.000 |
|  | Item rep. | 0.01 | [−0.062, 0.081] | 1.000 | 0.622 |
|  | Language × Place | 0.028 | [−0.055, 0.108] | 0.984 | 0.752 |
| COG | Intercept | −0.386 | [−0.520, -0.250] | 0.000 | 1.000 |
|  | Language | 0.303 | [0.196, 0.411] | 0.000 | 1.000 |
|  | Place | 0.647 | [0.522, 0.771] | 0.000 | 1.000 |
|  | F1 | 0.042 | [−0.005, 0.092] | 1.000 | 0.953 |
|  | F2 | −0.037 | [−0.083, 0.011] | 1.000 | 0.942 |
|  | Item rep. | 0.001 | [−0.049, 0.049] | 1.000 | 0.520 |
|  | Language × Place | 0.291 | [0.178, 0.396] | 0.000 | 1.000 |
| Kurtosis | Intercept | 0.433 | [0.275, 0.589] | 0.000 | 1.000 |
|  | Language | −0.233 | [−0.301, -0.163] | 0.000 | 1.000 |
|  | Place | −0.829 | [−0.969, -0.677] | 0.000 | 1.000 |
|  | F1 | 0.003 | [−0.040, 0.048] | 1.000 | 0.552 |
|  | F2 | 0.009 | [−0.042, 0.056] | 1.000 | 0.656 |
|  | Item rep. | −0.044 | [−0.092, 0.004] | 1.000 | 0.963 |
|  | Language × Place | −0.191 | [−0.259, -0.123] | 0.000 | 1.000 |
| SD | Intercept | −0.501 | [−0.633, -0.370] | 0.000 | 1.000 |
|  | Language | 0.234 | [0.145, 0.321] | 0.000 | 1.000 |
|  | Place | 0.782 | [0.661, 0.900] | 0.000 | 1.000 |
|  | F1 | 0.011 | [−0.034, 0.058] | 1.000 | 0.687 |
|  | F2 | −0.017 | [−0.066, 0.029] | 1.000 | 0.765 |
|  | Item rep. | 0.024 | [−0.026, 0.071] | 1.000 | 0.843 |
|  | Language × Place | 0.204 | [0.118, 0.291] | 0.000 | 1.000 |
| Skewness | Intercept | 0.546 | [0.350, 0.745] | 0.000 | 1.000 |
|  | Language | −0.126 | [−0.228, -0.027] | 0.286 | 0.992 |
|  | Place | −0.865 | [−1.042, -0.680] | 0.000 | 1.000 |
|  | F1 | 0.006 | [−0.059, 0.076] | 1.000 | 0.555 |
|  | F2 | 0.014 | [−0.030, 0.059] | 1.000 | 0.732 |
|  | Item rep. | −0.013 | [−0.066, 0.046] | 1.000 | 0.684 |
|  | Language × Place | −0.092 | [−0.192, 0.007] | 0.561 | 0.967 |

Appendix E

## Bayesian data analysis

This study employs Bayesian Data Analysis for quantitative inferential statistics. Specifically, this implies that we use Bayesian *credible intervals*—and other metrics—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 β, but rather we draw a sample of 4,000 plausible values for β. This allows us to quantify our uncertainty regarding β by summaryzing the distribution of those values. We will use 4 statistics to describe the posterior distribution: (1) the mean, (2) the highest density credible interval (HDI), (3) the proportion of the HDI that falls within a Region of Practical Equivalence (ROPE), and (4) the Maximum Probability of Effect (MPE). The mean provides a point estimate for the distribution. The 95% highest density credible interval provides bounds for the effect. The ROPE designates a region of practical equivalence for a negligible effect and calculates the proportion of the HDI that falls within this interval.[[2]](#footnote-61) The MPE calculates the proportion of the posterior distribution that is of the median’s sign (or the probability that the effect is positive or negative).

If, for instance, a hypothesis states that β > 0, we judge there to be *compelling evidence* for this hypothesis if the mean point estimate is a positive number, if the 95% credible interval of β does not contain 0 and is outside the ROPE by a reasonably clear margin, and the posterior *P*(β > 0) is close to one. Together these four statistics allow us to quantify our uncertainty and provide an intuitive interpretation of any given effect. Consider a case in which the posterior mean of β is 100 and the 95% credible interval is [40, 160]. The interval tells us that we can be 95% certain the *true* value of β 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.

Appendix F

## About this document

This document was written in RMarkdown using papaja (Aust & Barth, 2018) and serves as a project report for our research group. 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.

## Session info

> setting value   
> version R version 3.6.0 (2019-04-26)  
> os macOS Mojave 10.14.6   
> system x86\_64, darwin15.6.0   
> ui X11   
> language (EN)   
> collate en\_US.UTF-8   
> ctype en\_US.UTF-8   
> tz America/New\_York   
> date 2019-10-02

> loadedversion date  
> abind 1.4-5 2016-07-21  
> academicWriteR 0.4.0 2019-09-29  
> arrayhelpers 1.0-20160527 2016-05-28  
> assertthat 0.2.1 2019-03-21  
> backports 1.1.4 2019-04-10  
> base64enc 0.1-3 2015-07-28  
> bayesplot 1.7.0 2019-05-23  
> bayestestR 0.3.0 2019-09-22  
> beeswarm 0.2.3 2016-04-25  
> bookdown 0.13 2019-08-21  
> boot 1.3-23 2019-07-05  
> bridgesampling 0.7-2 2019-07-21  
> brms 2.10.0 2019-08-29  
> Brobdingnag 1.2-6 2018-08-13  
> broom 0.5.2 2019-04-07  
> broom.mixed 0.2.4 2019-02-21  
> callr 3.3.2 2019-09-22  
> cli 1.1.0 2019-03-19  
> coda 0.19-3 2019-07-05  
> codetools 0.2-16 2018-12-24  
> colorspace 1.4-1 2019-03-18  
> colourpicker 1.0 2017-09-27  
> crayon 1.3.4 2017-09-16  
> crosstalk 1.0.0 2016-12-21  
> data.table 1.12.2 2019-04-07  
> desc 1.2.0 2018-05-01  
> devtools 2.2.1 2019-09-24  
> digest 0.6.21 2019-09-20  
> dplyr 0.8.3 2019-07-04  
> DT 0.9 2019-09-17  
> dygraphs 1.1.1.6 2018-07-11  
> ellipsis 0.3.0 2019-09-20  
> evaluate 0.14 2019-05-28  
> flextable 0.5.6.002 2019-09-29  
> forcats 0.4.0 2019-02-17  
> fs 1.3.1 2019-05-06  
> future 1.14.0 2019-07-02  
> gdtools 0.2.0 2019-09-03  
> generics 0.0.2 2018-11-29  
> ggbeeswarm 0.6.0 2017-08-07  
> ggplot2 3.2.1 2019-08-10  
> ggridges 0.5.1 2018-09-27  
> ggstance 0.3.3 2019-08-19  
> globals 0.12.4 2018-10-11  
> glue 1.3.1 2019-03-12  
> gridExtra 2.3 2017-09-09  
> gtable 0.3.0 2019-03-25  
> gtools 3.8.1 2018-06-26  
> here 0.1 2017-05-28  
> highr 0.8 2019-03-20  
> hms 0.5.1 2019-08-23  
> htmltools 0.3.6 2017-04-28  
> htmlwidgets 1.3 2018-09-30  
> httpuv 1.5.2 2019-09-11  
> igraph 1.2.4.1 2019-04-22  
> inline 0.3.15 2018-05-18  
> insight 0.5.0 2019-09-07  
> knitr 1.25 2019-09-18  
> later 0.8.0 2019-02-11  
> lattice 0.20-38 2018-11-04  
> lazyeval 0.2.2 2019-03-15  
> lifecycle 0.1.0 2019-08-01  
> listenv 0.7.0 2018-01-21  
> lme4 1.1-21 2019-03-05  
> loo 2.1.0 2019-03-13  
> magrittr 1.5 2014-11-22  
> markdown 1.1 2019-08-07  
> MASS 7.3-51.4 2019-03-31  
> Matrix 1.2-17 2019-03-22  
> matrixStats 0.55.0 2019-09-07  
> memoise 1.1.0 2017-04-21  
> mime 0.7 2019-06-11  
> miniUI 0.1.1.1 2018-05-18  
> minqa 1.2.4 2014-10-09  
> modelr 0.1.5 2019-08-08  
> munsell 0.5.0 2018-06-12  
> nlme 3.1-141 2019-08-01  
> nloptr 1.2.1 2018-10-03  
> officer 0.3.5 2019-06-23  
> papaja 0.1.0.9842 2019-06-03  
> pillar 1.4.2 2019-06-29  
> pkgbuild 1.0.5 2019-08-26  
> pkgconfig 2.0.3 2019-09-22  
> pkgload 1.0.2 2018-10-29  
> plyr 1.8.4 2016-06-08  
> prettyunits 1.0.2 2015-07-13  
> processx 3.4.1 2019-07-18  
> promises 1.0.1 2018-04-13  
> ps 1.3.0 2018-12-21  
> purrr 0.3.2 2019-03-15  
> R6 2.4.0 2019-02-14  
> Rcpp 1.0.2 2019-07-25  
> readr 1.3.1 2018-12-21  
> remotes 2.1.0 2019-06-24  
> reshape2 1.4.3 2017-12-11  
> rlang 0.4.0 2019-06-25  
> rmarkdown 1.15 2019-08-21  
> rprojroot 1.3-2 2018-01-03  
> rsconnect 0.8.15 2019-07-22  
> rstan 2.19.2 2019-07-09  
> rstantools 2.0.0 2019-09-15  
> scales 1.0.0 2018-08-09  
> sessioninfo 1.1.1 2018-11-05  
> shiny 1.3.2 2019-04-22  
> shinyjs 1.0 2018-01-08  
> shinystan 2.5.0 2018-05-01  
> shinythemes 1.1.2 2018-11-06  
> StanHeaders 2.19.0 2019-09-07  
> stringi 1.4.3 2019-03-12  
> stringr 1.4.0 2019-02-10  
> svUnit 0.7-12 2014-03-02  
> systemfonts 0.1.1 2019-07-01  
> testthat 2.2.1 2019-07-25  
> threejs 0.3.1 2017-08-13  
> tibble 2.1.3 2019-06-06  
> tidybayes 1.1.0.9000 2019-08-27  
> tidyr 1.0.0 2019-09-11  
> tidyselect 0.2.5 2018-10-11  
> TMB 1.7.15 2018-11-09  
> usethis 1.5.1 2019-07-04  
> uuid 0.1-2 2015-07-28  
> vctrs 0.2.0 2019-07-05  
> vipor 0.4.5 2017-03-22  
> withr 2.1.2 2018-03-15  
> xfun 0.9 2019-08-21  
> xml2 1.2.2 2019-08-09  
> xtable 1.8-4 2019-04-21  
> xts 0.11-2 2018-11-05  
> yaml 2.2.0 2018-07-25  
> zeallot 0.1.0 2018-01-28  
> zip 2.0.4 2019-09-01  
> zoo 1.8-6 2019-05-28

1. See Schoot and Depaoli (2014) and Vasishth, Nicenboim, Beckman, Li, and Kong (2018) for tutorials and in depth explanations related to BDA in the psychological and speech sciences. [↑](#footnote-ref-27)
2. We utilize a ROPE of ± 1 for standardized values. For non-standardized values Kruschke (2018) recommomends using the formula in (1) [↑](#footnote-ref-61)