nature human behaviour

Article https://doi.org/10.1038/s41562-022-01452-1

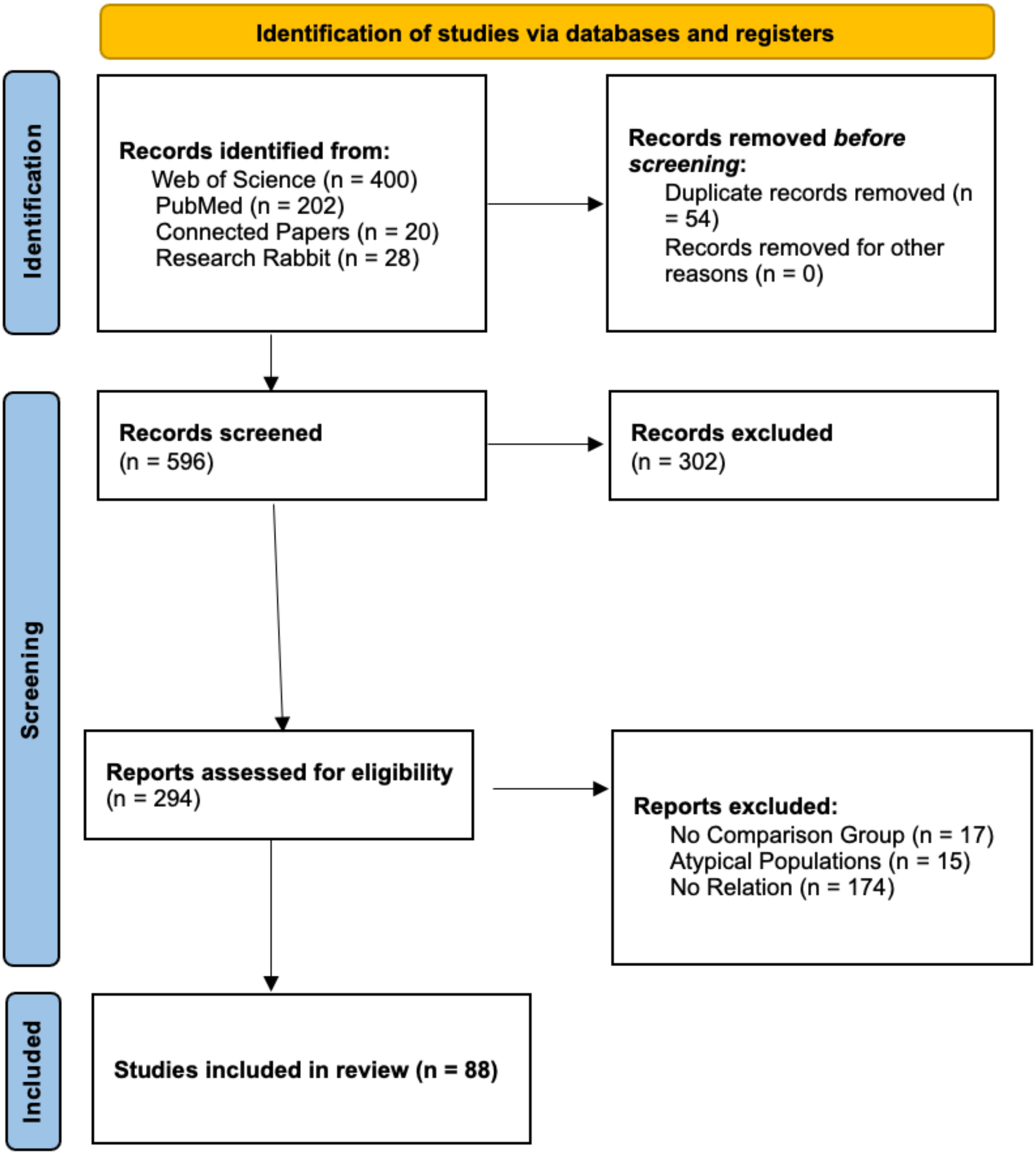
A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech

In the format provided by the authors and unedited

Supplementary Information

Table of Contents

S1: Details about the Systematic Search



Supplementary Figure 1.1: PRISMA chart of the systematic review process

Topic	No.	Item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Introduction
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Methods
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Methods
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Methods
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Methods
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Methods
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Methods
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Methods

Topic	No.	Item	Location where item is reported
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Methods
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Methods
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	Methods
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Methods
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Methods
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Methods
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Methods
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Methods
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Methods
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Methods
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	S1.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	NA

Topic	No.	Item	Location where item is reported
Study characteristics	17	Cite each included study and present its characteristics.	S2
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	S1.1
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Results
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Results
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Results
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Results
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Results
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Methods
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Results
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Discussion
	23b	Discuss any limitations of the evidence included in the review.	Discussion
	23c	Discuss any limitations of the review processes used.	Discussion
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	NA
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Methods

Topic	No.	Item	Location where item is reported
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	NA
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Acknowledgements
Competing interests	26	Declare any competing interests of review authors.	Competing Interests Statement
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Data & Code availability statement

Supplementary Table 1.2: PRISMA Main Checklist

Topic	No.	Item	Reported?
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesize results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	Yes
Registration	12	Provide the register name and registration number.	NA

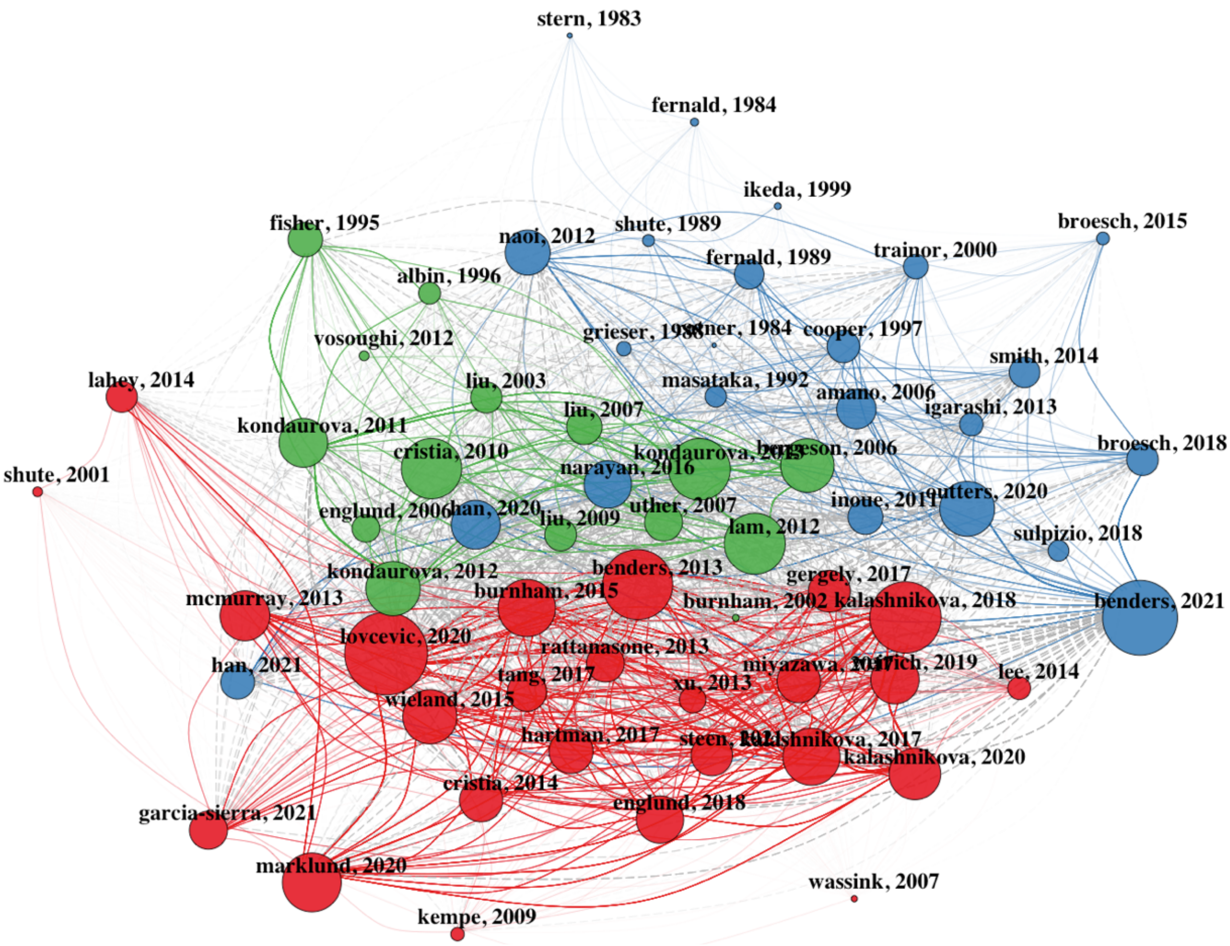
Supplementary Table 1.3: PRISMA Abstract Checklist

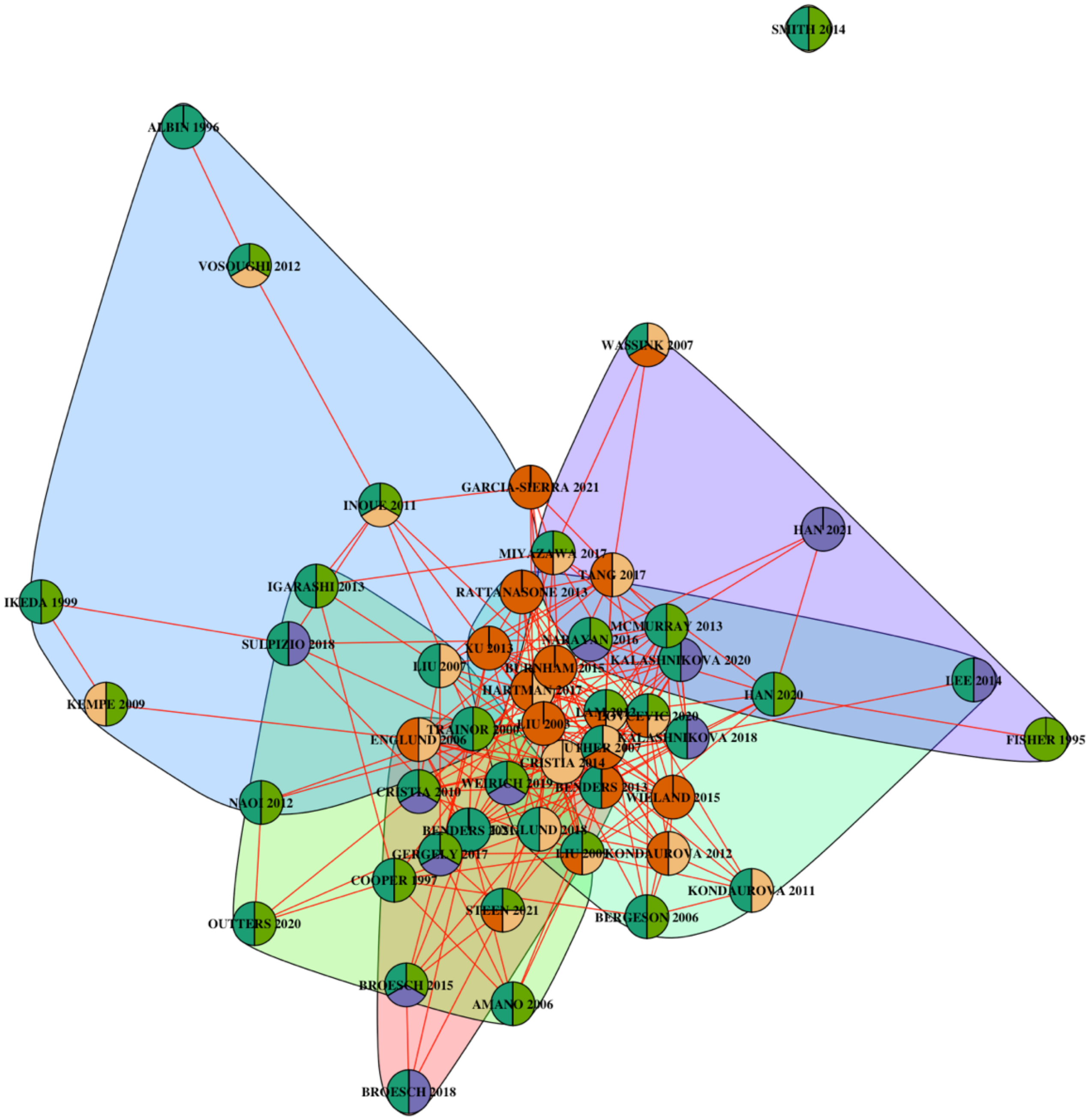
S1.1: Risk of Bias Assessment

Despite being less prone to bias than more subjective literature overviews, systematic searches and meta-analyses cannot completely avoid bias. In this section, we discuss some of the potential biases in our systematic search and selection process. Firstly, *our choice of search terms may select a biased subset of the literature*. In order to counteract this potential source of bias, we aimed to make our list of search terms as inclusive as possible. We conducted initial searches after carefully reading through relevant papers and included additional terms before conducting the final systematic search. We also performed forward and backward searches of the literature using cutting-edge bibliography tools (Research Rabbit and Connected Papers) to expand the scope of our search to relevant studies that were not found initially. Secondly, the *published literature itself may represent a biased subset of the literature available* on the acoustic features of IDS, to a greater extent reporting outcome measures for which manipulation created a significant effect. In order to counteract the effects of this publication bias in the meta-analysis, we carried out the following. i) In the literature search we included both published and grey literature, such as pre-prints, conference proceedings, etc. and informally solicited literature suggestions on twitter and from experts. ii) In the manuscript, we actively encourage researchers with unpublished and published work to submit their experimental results to an open repository with the data from our meta-analysis (MetaLab: https://langcog.github.io/metalab/). iii) In the meta-analysis, we assess the extent to which the meta-analytic estimates change under different assumptions of publication bias in the literature, by conducting quantitative sensitivity analyses. We should also note that a related source of bias may manifest itself in *the study selection process where authors exclude papers not conforming to their hypotheses*. However, because this project concerns estimates of acoustic features above and beyond statistical significance tests, we had no specific directional hypotheses to test and strong incentives to include as much data as possible. Thirdly, *bias might arise as a function of the reporting of estimates* (e.g., studies with missing estimates of uncertainty may be of systematically lower quality than other studies). Because most of the papers with missing data were older papers, we chose not to contact the original authors to provide the missing data because we know from previous work that answers are extremely unlikely. We instead decided to impute the missing measures of uncertainty, as outlined in Section 2.1. This imputation process was shown not to bias any of the results, as shown in Section S3, and better counteracts this potential source of bias than simply excluding studies not reporting measures of uncertainty. In general, although no analysis can remain completely unbiased, we hope this project can serve as a first step towards a cumulative self-correcting enterprise. Accordingly, we make our data openly available as a Community Augmented Meta-Analysis on the MetaLab website[151]. This makes it possible to critique, integrate and update our selection of studies in a straightforward manner.

S2: Citation Networks

The upper network shows the co-citation coupling strength (i.e., the number of times two studies are cited together by a new article as well as bibliographic similarity) for only the journal articles of the final sample of cited studies. The colour and thickness of the lines represent clusters of strong citation links. The lower direct-citation network shows which studies cite each other. The colours represent clusters of strong direct-citation links. In the below direct-citation plot, the colours of the nodes represent the acoustic measures under investigation; specifically, dark green is f_0, light green is f_0 variability, orange is vowel space area, purple is articulation rate, and light orange is vowel duration.





Supplementary Figure 2.1: Coupling (upper) and Direct-Citation (lower) Networks of Studies on IDS.

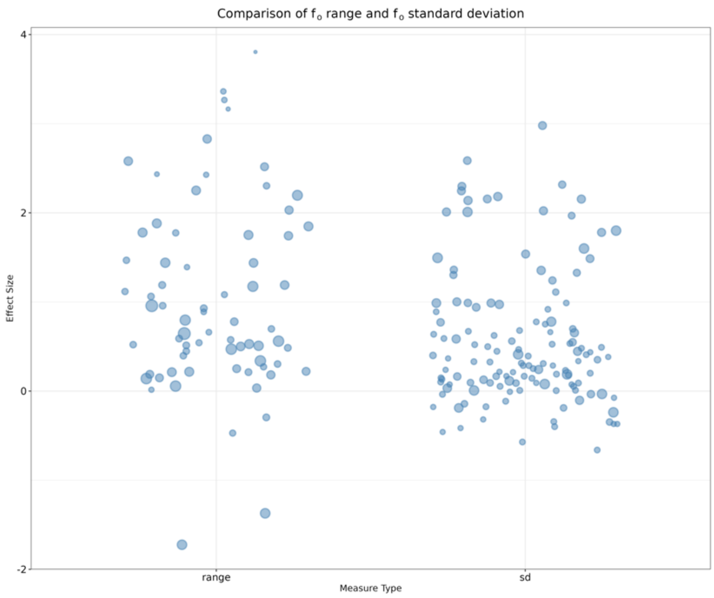
S3: Imputation Process

In order to incorporate the statistical uncertainty associated with the partially stochastic nature of this imputation process (Azur et al., 2011; Sterne et al., 2009), we constructed 20 datasets with sample size, mean values for each acoustic variable, and existing standard deviation values as predictors. The standard deviation values of the imputed datasets were checked for similarity to the reported standard deviations and post-processed to include only values within the range of the existing standard deviation values. In order to check that this process of multiple imputation did not bias the estimation of the overall effect size for each acoustic measure, we compared the estimates of the intercepts-only models for the imputed and non-imputed datasets, as shown in Supplementary Table 3.1 below. There does not appear to be evidence of bias, as the effect size estimate of the models with the imputed datasets lies within the credible interval of the non-imputed datasets in each case.

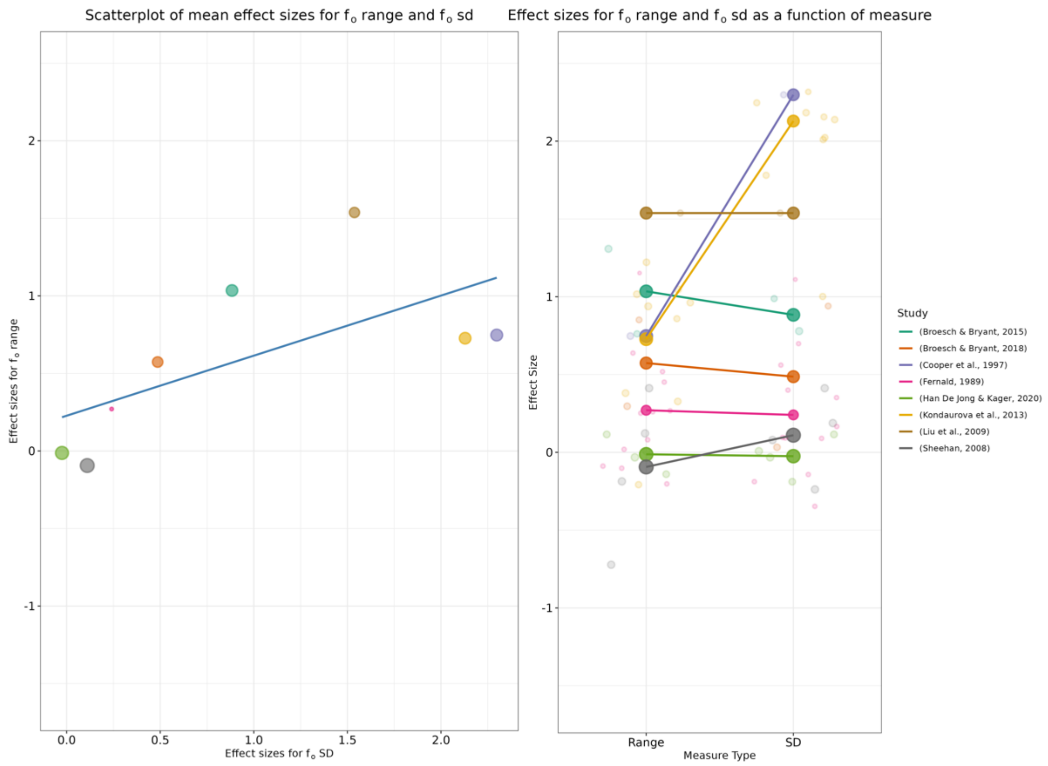
Acoustic Measure	**Intercept Estimate *Without* Imputation (n = total observations)**	**Effect Size Estimate *With* Imputation (n = total observations)**
f_0	1.09 [0.83; 1.34] (n = 250)	1.17 [0.86; 1.45] (n = 262)
f_0 **Variability**	0.76 [0.49; 1.00] (n = 208)	0.69 [0.44; 1.92] (n = 223)
Vowel Space Area	0.49 [-0.08; 1.09] (n = 51)	0.66 [0.34; 0.98] (n = 107)
Articulation Rate	-0.91 [-1.42; -0.42] (n = 56)	-1.05 [-1.53; -0.60] (n = 60)
Vowel Duration	0.47 [0.02; 0.91] (n = 72)	0.48 [0.08; 0.88] (n = 82)

Supplementary Table 3.1: An overview of the extent to which imputation has influenced the overall estimation of effect sizes for each acoustic measure

S4: Comparison of f_0 range and f_0 standard deviation



Supplementary Figure 4.1: A plot showing the distribution of effect sizes for fo range and fo standard deviation. The similar distributions speak in favor of our choice to combine the measures into one measure of f_0 variability.



Supplementary Figure 4.2: A correlation scatter plot showing the distribution of average effect sizes for the studies reporting both measures (left) and a plot of the effect sizes as a function of measure (right). A Bayesian multivariate model with range and standard deviation as separate outcomes shows a strong correlation between the two measures 0.73 [0.38; 0.98] (without Kondaurova et al., 2013), as these authors report f_0 range in semitones and f_0 standard deviation in Hz). This estimate is based on a total of 573 participants across the 8 studies.

S5: Choice of Priors, Prior and Posterior Predictive Checks, Prior-Posterior Update Plots, Prior Robustness Checks

S5.1: Choice of Priors

We chose weakly informative priors in order to ensure that their influence on the meta-analytic estimates was small and to discount extreme effect sizes as unlikely[152, 153]

(cf. Lemoine, 2019; Gelman, Simpson & Betancourt, 2017); for the overall effect size, we chose a Gaussian distribution with a mean of 0 and standard deviation of 2.5 based on our prior expectations for effect sizes. This prior implies that we expect approximately 95% of the effect size distribution to be between -5 and 5. For the slope of the model, we encoded our expectations with a Gaussian prior with a mean of 0 and a standard deviation of 1, which implies that we expect the vast majority of values for the coefficient of the effect size difference between ADS and IDS to be between -2 and 2. For the heterogeneity of the effects (i.e., the standard deviation of random effects), we chose a positive truncated normal distribution with a mean of 1 and standard deviation of 1. For the degrees of freedom parameter, ν, of the Student's t-distribution, a gamma distribution with a shape parameter of 2 and a scale parameter of 0.1 was chosen. This ensures that the model remains robust to the influence of outliers. Prior predictive checks were performed to ensure that model predictions for plausible values of effect sizes would only exclude implausibly high or low values on the basis of the priors[152].

The models were fitted with Hamiltonian Monte Carlo samplers with 2 parallel chains with 5,000 iterations each, an adapt delta of 0.99 and a maximum tree depth of 20 in order to ensure no divergence in the estimation process. The quality of the models was assessed by i) ensuring Rhat statistics to be lower than 1.1, ii) carrying out prior and posterior predictive checks, iii) plotting prior against posterior estimates and assessing whether the posteriors had lower variance than the priors, iv) ensuring no divergences in the process of estimation, v) checking that the number of effective bulk and tail samples was above 200, vi) conducting prior sensitivity analyses.

For the intercepts-only and full models, we used the following *brms*[142] formula and priors, with a student t likelihood for all of the effect sizes measures, as shown below:

Intercepts Model Structure: *Effect_Size | se(Effect_Size_se) ~ 1 + (1 | Lang/StudySite/measure)*

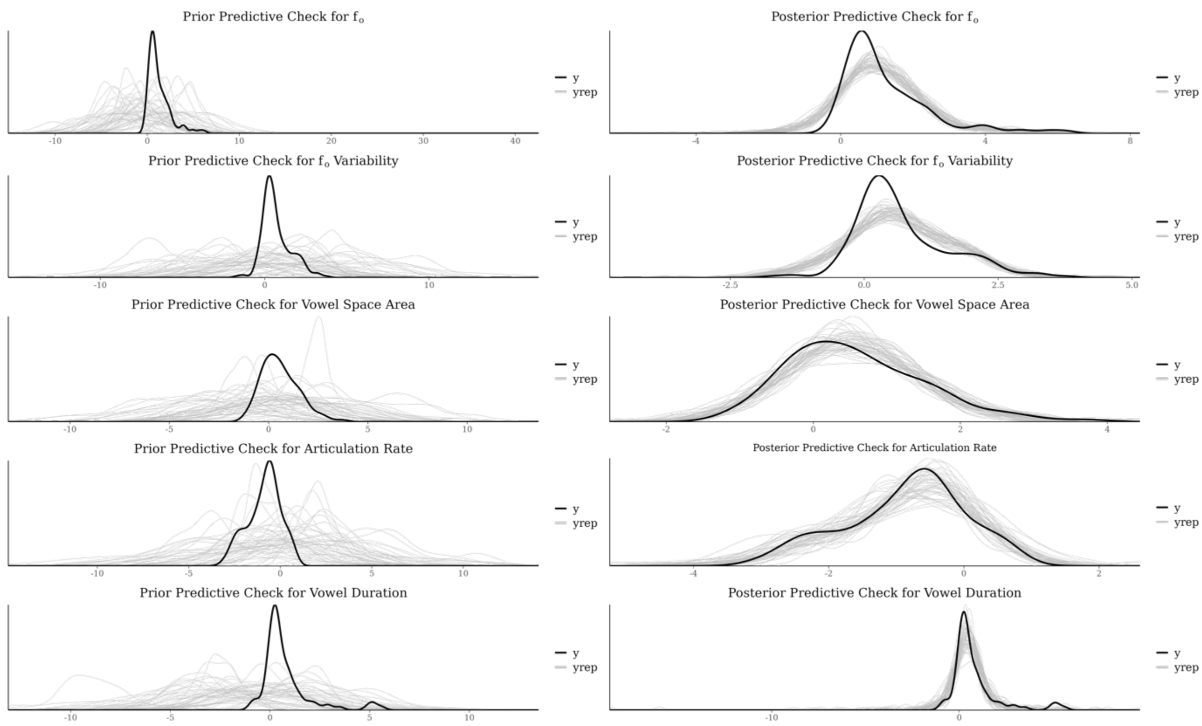
Full Model Structure: *Effect_Size | se(Effect_Size_se) ~ 1 + Age + Lang + Environment + Task + (1 | Lang/StudySite/measure))*

Models	Intercepts	Slopes	SD	DoF
Intercepts Model	N(0,2.5)	-	N(1,1)	G(2,.1)
Full Moderators Model	N(0,2.5)	N(0,1) for Task & Environment N(0.05) for Age	N(1,1)	G(2,.1)

Supplementary Table 5.1: Priors for the parameters in the intercepts-only model and full model with all moderators. N() refers to a normal distribution, G() indicates a gamma distribution, lkj() refers to the Lewandowski-Kurowicka-Joe distribution. DoF refers to Degrees of Freedom parameter (or ν).

S5.2: Prior & Posterior Predictive Checks

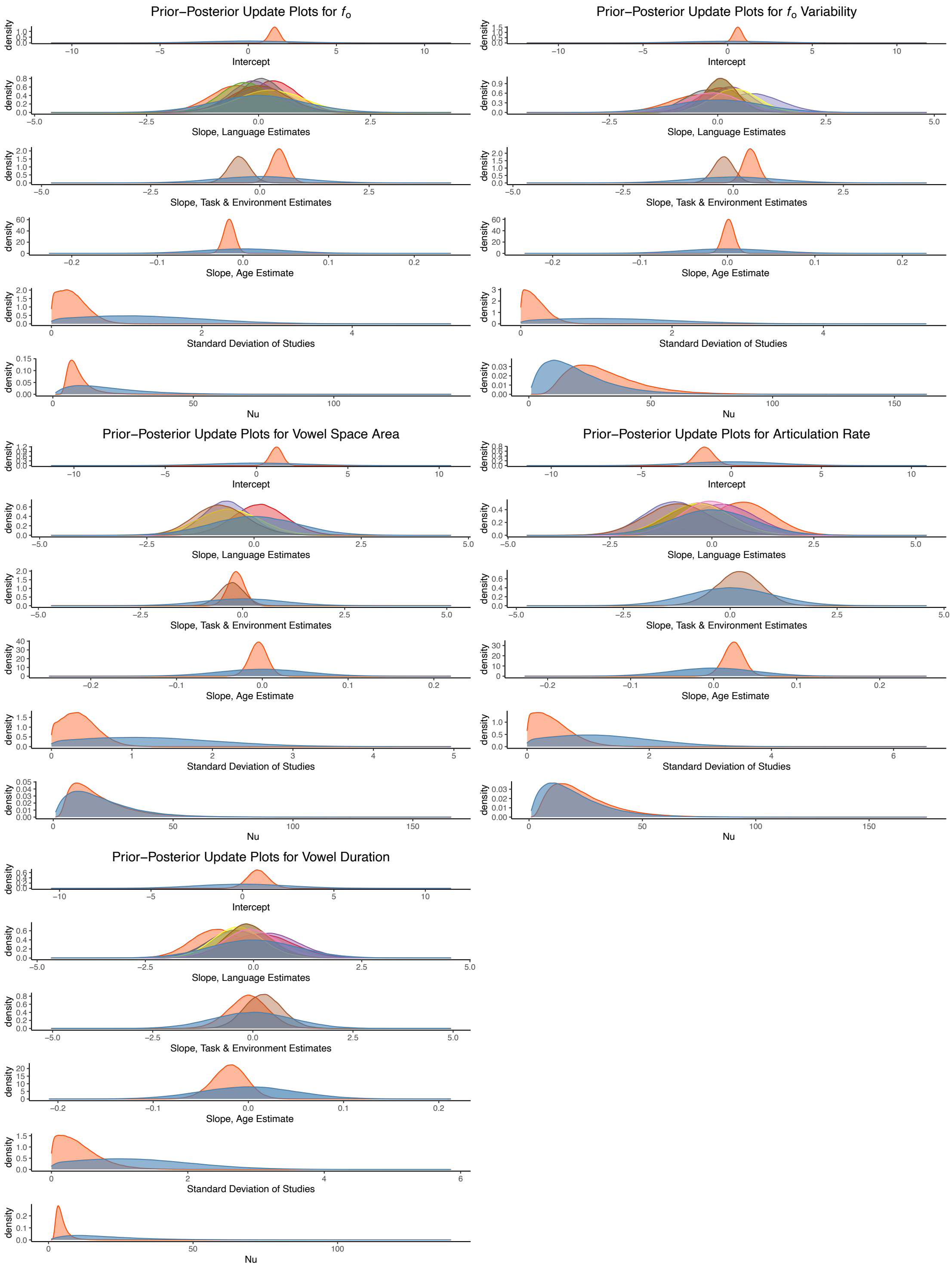
As noted above, we performed quality checks of the models by carrying out prior and posterior predictive checks. The below prior predictive checks (on the left) indicate that our priors predict values within the order of magnitude of the distribution. The posterior predictive checks (on the right) indicate that the models have captured the distributions of data for each of the acoustic measures. These plots provide reassurance that our models capture relevant aspects of the overall distributions of dependent variables.



Supplementary Figure 5.2.1: Plot of the prior and posterior predictive checks (grey) and observed meta-analytic data (black) for the acoustic measures.

S5.3: Prior-Posterior Update Plots

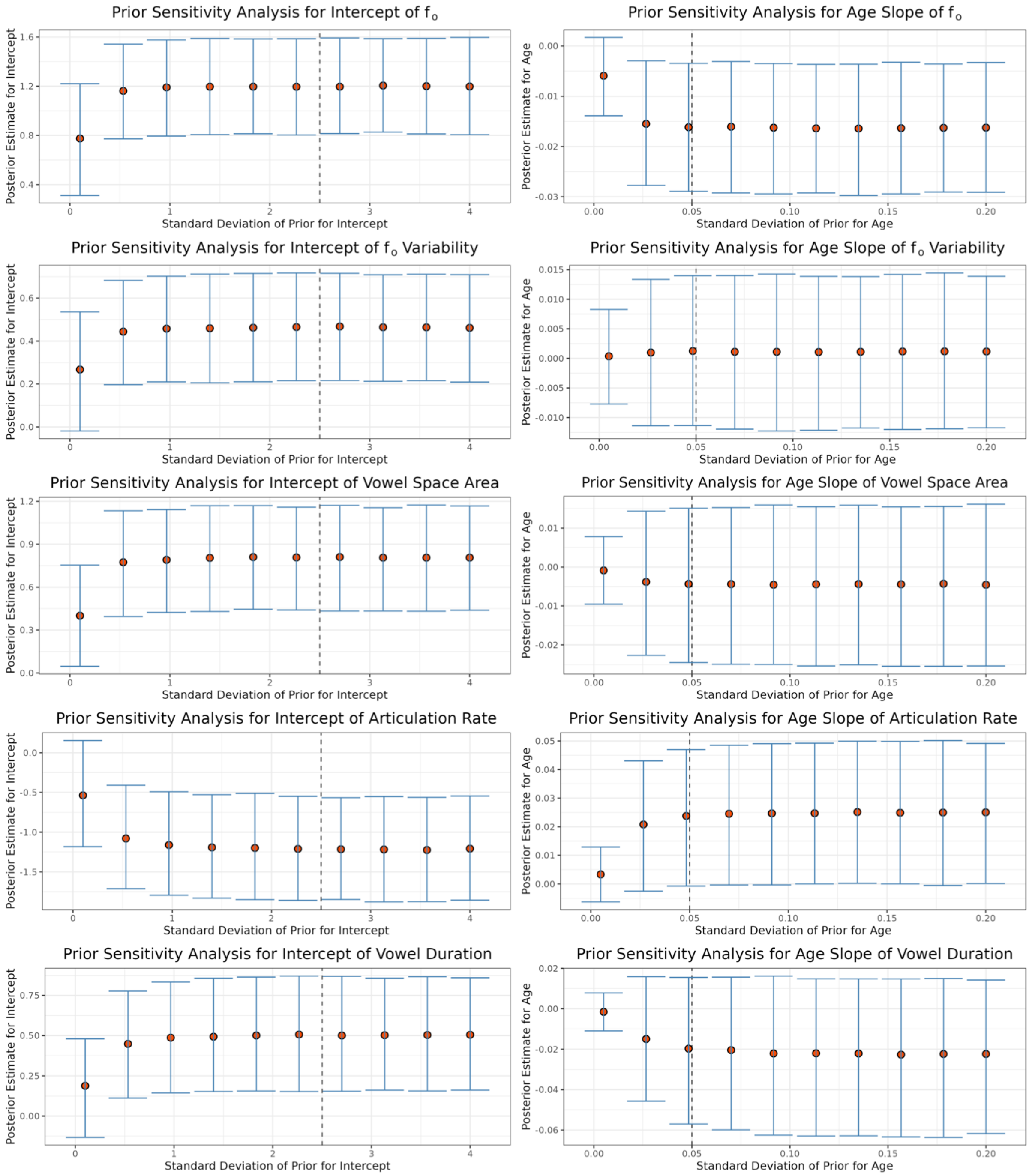
A second quality check of the models was carried out by plotting the prior distributions against the posterior estimates of the model. As shown in the below plots, the posteriors exhibit lower variance than the priors. These plots thus indicate that the models have learned from the data and provide additional reassurance that our models have captured relevant information.



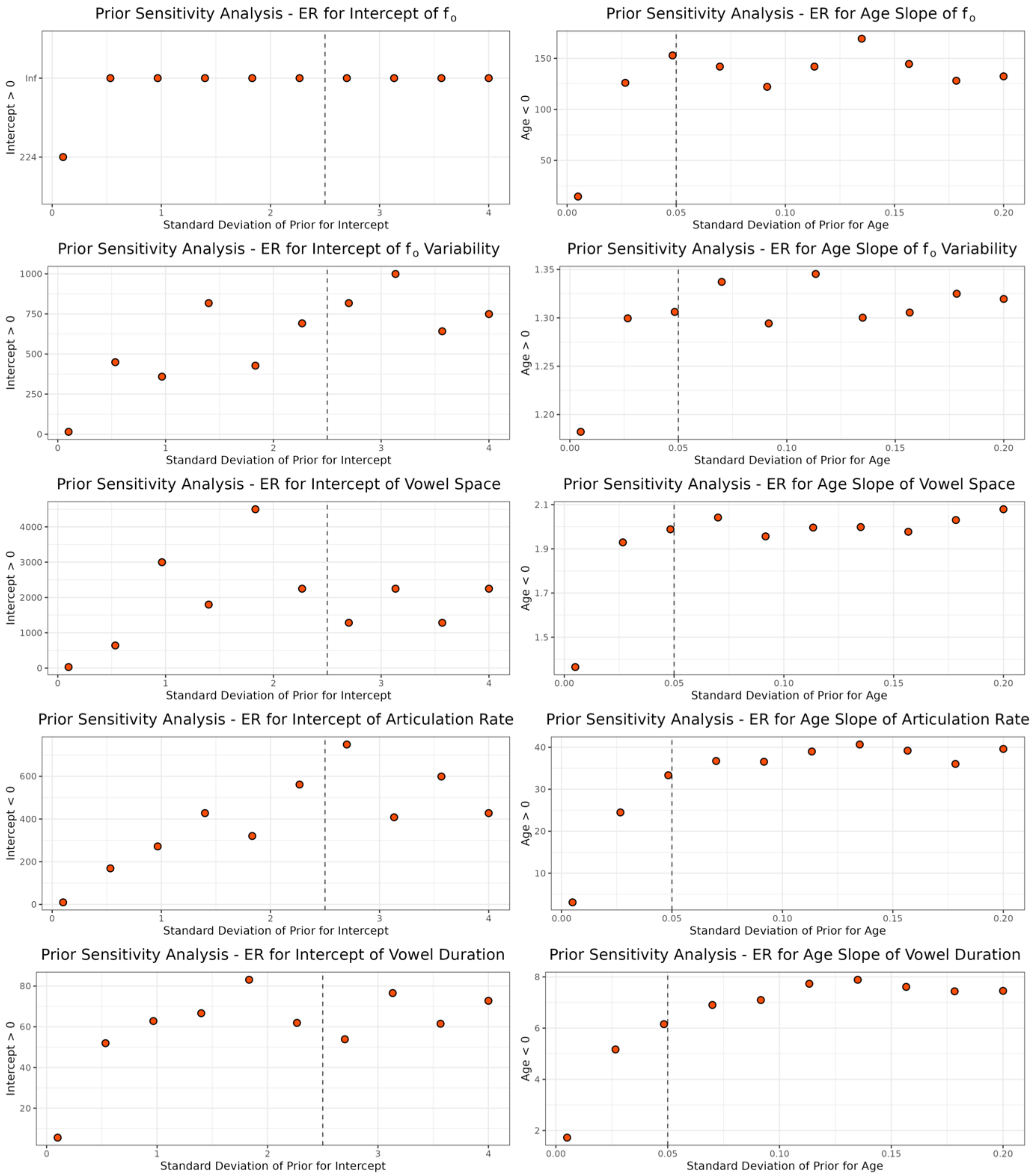
Supplementary Figure 5.3.1: Panel of prior-posterior update plots for the intercept, slope, standard deviation, and nu for each of the acoustic variables under investigation. The prior distributions are represented in blue. In the plots of task and environment, task is represented by orange and environment is represented by brown.

S5.4: Prior Sensitivity Analysis for Intercept & Slope

A third quality check of the models was performed by assessing the extent to which the uncertainty of our priors affected posterior estimates. Because the posterior estimates (on the y-axis) exhibit stability at our choices of priors (i.e., the dashed vertical line), these plots provide reassurance that our choice of priors did not unduly affect model estimations.

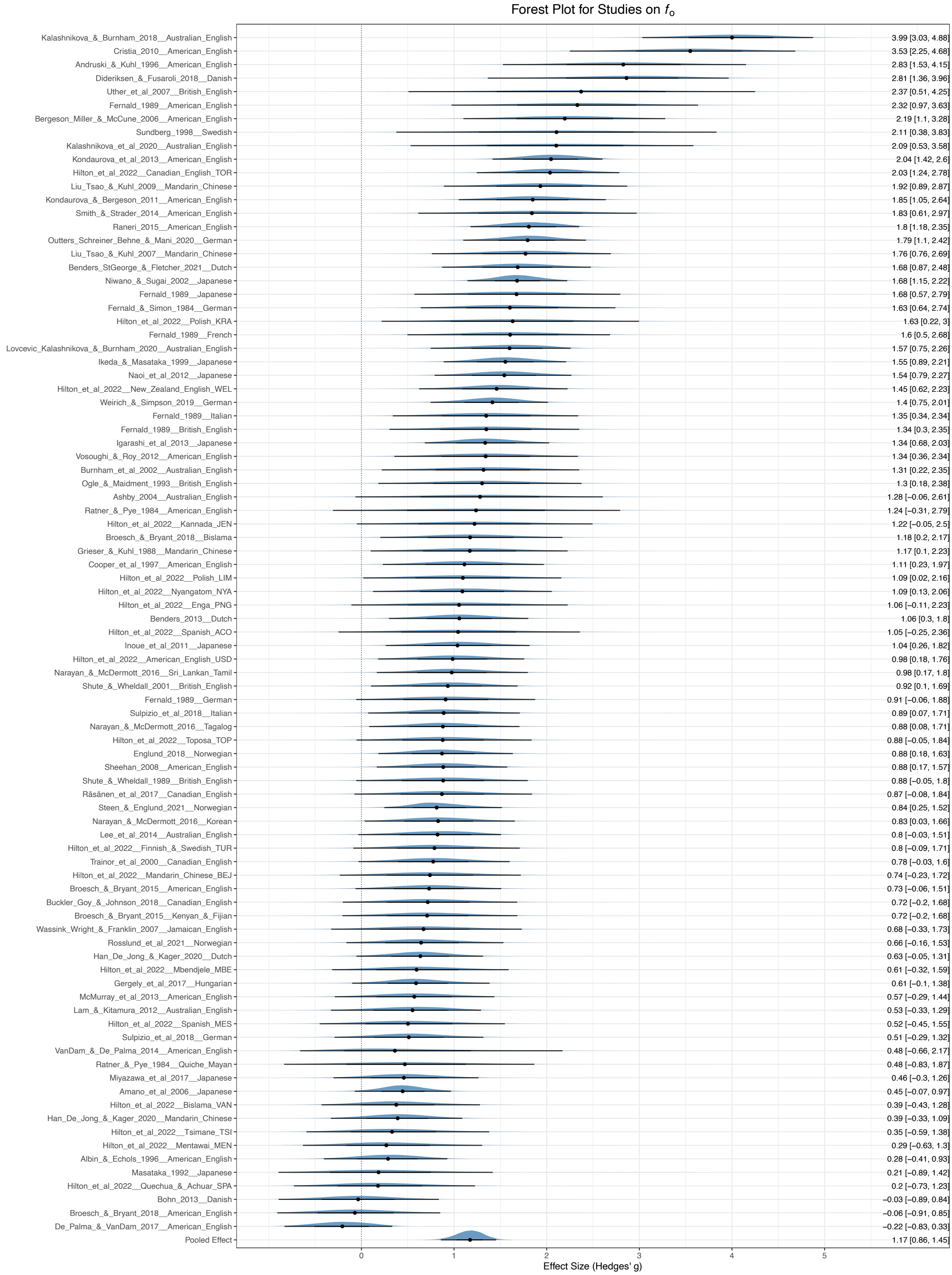


Supplementary Figure 5.4.1: Panel of plots showing how the intercept and age estimates for each acoustic variable change with different standard deviations for the priors. The vertical dashed line indicates the standard deviation of the prior chosen for the models. The centres of the error bars (orange points) indicate the posterior estimates for the intercept (left column) and age predictor (right column). The total sample sizes across studies for each of the estimates were 3401, 3006, 1702, 976, 1411 participants for f_o, f_o variability, vowel space area, articulation rate, vowel duration, respectively.

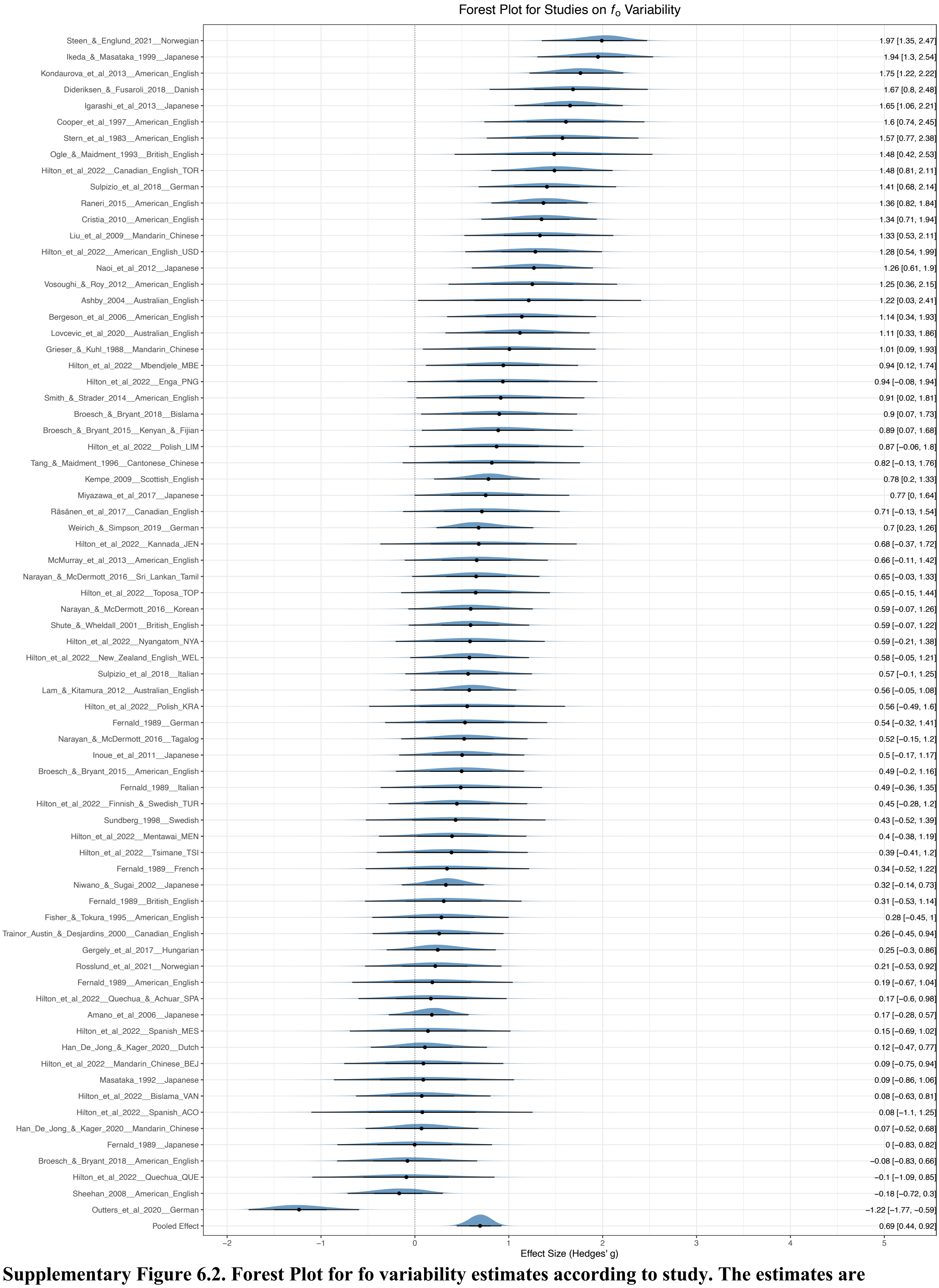


Supplementary Figure 5.4.2: Panel of plots showing how the evidence ratio (ER) for the intercept and age estimates for each acoustic variable change with different standard deviations for the priors. The vertical dashed line indicates the standard deviation of the prior chosen for the models.

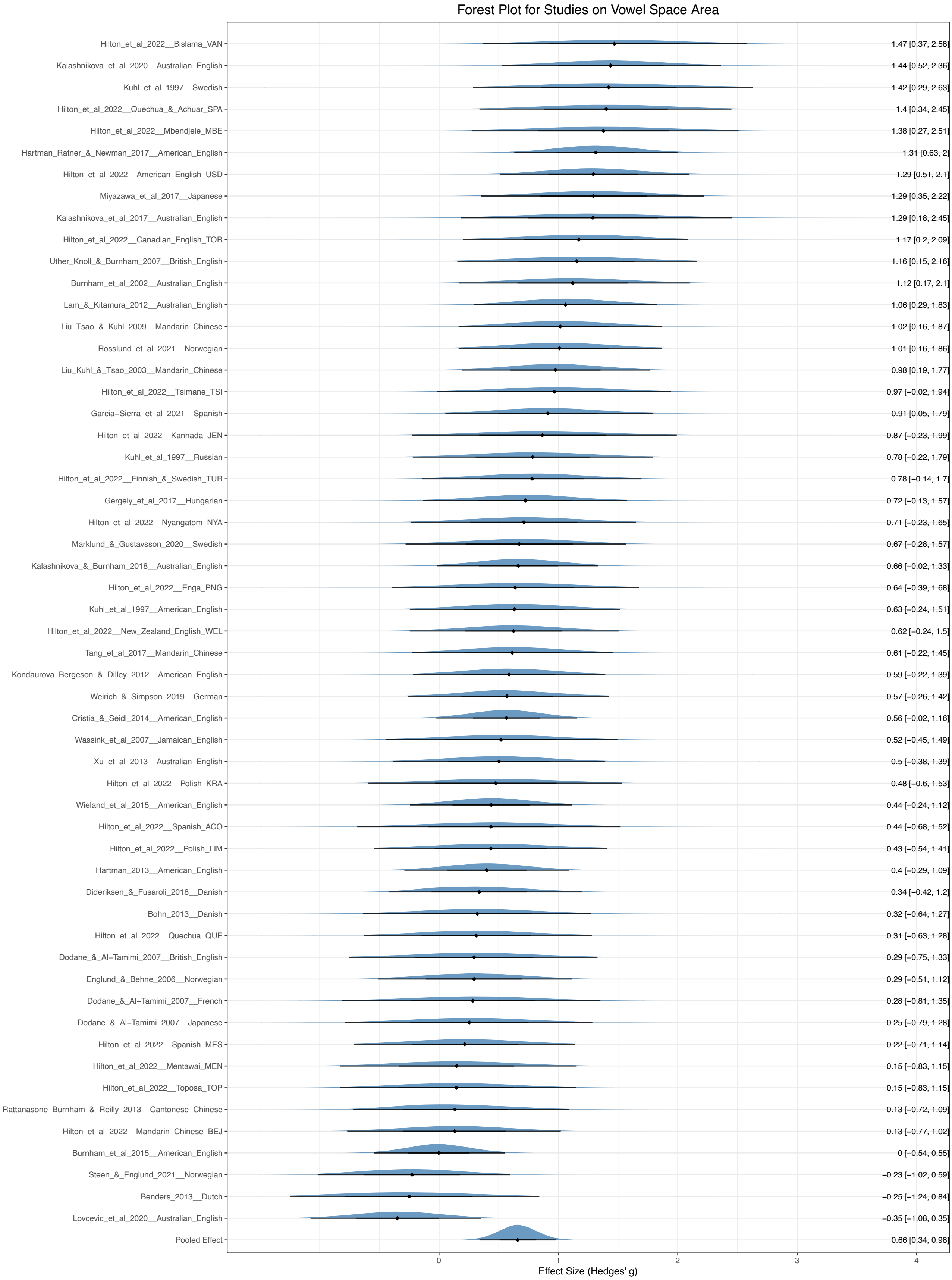
S6: Forest Plots & Overview Plots



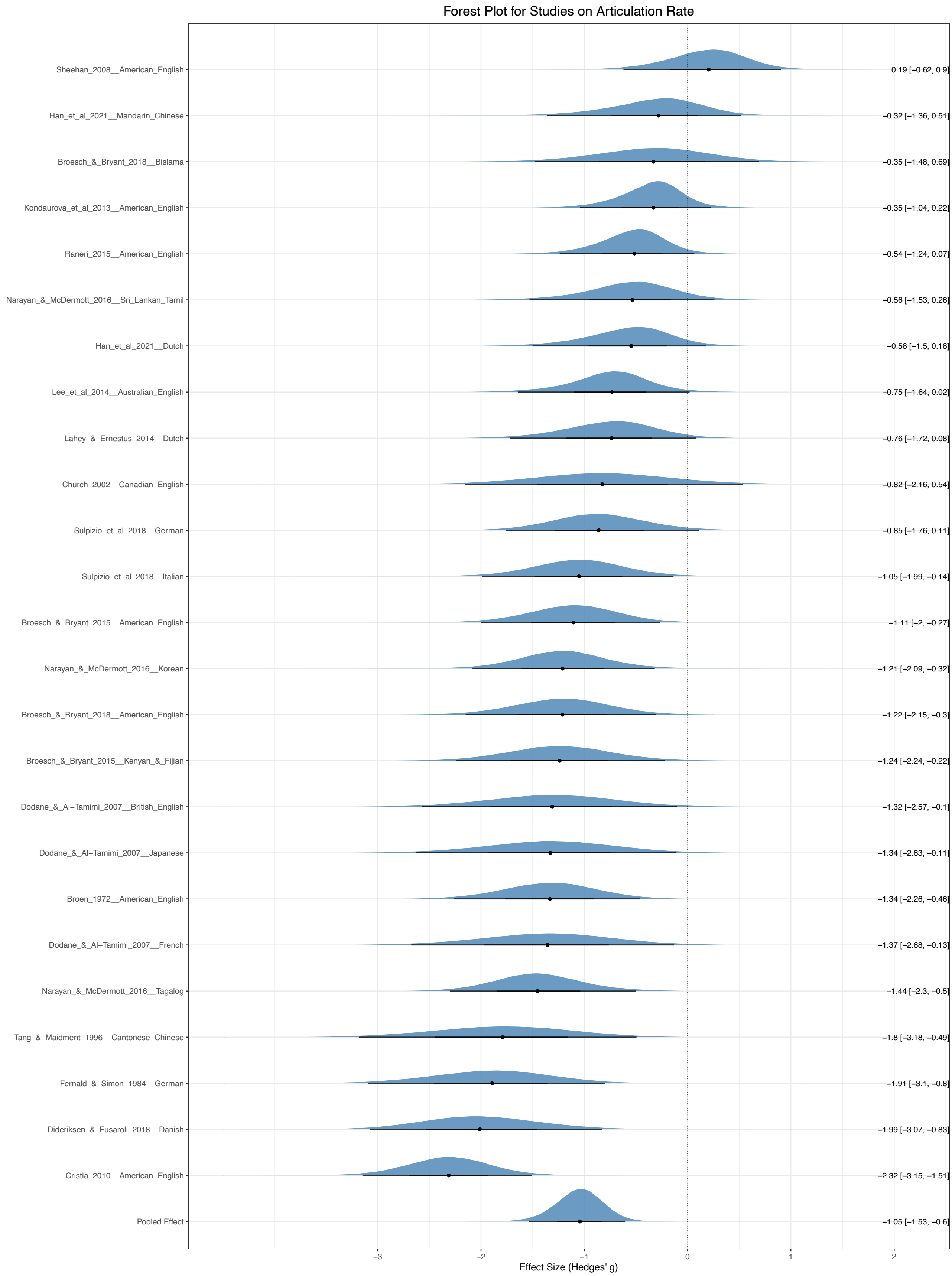
Supplementary Figure 6.1. Forest Plot for fo estimates according to study and language. The estimates are based on a total of 3401 participants across 60 studies investigating 33 distinct languages. The shaded areas indicate the posterior probability density of each estimate. The numbers to the right provide the estimated mean effect size (Hedges' *g*) and upper and lower 95% credible intervals). The estimates within each study are broken down according to language; this is especially evident for studies[1, 54], from which data on a diverse set of languages exist.



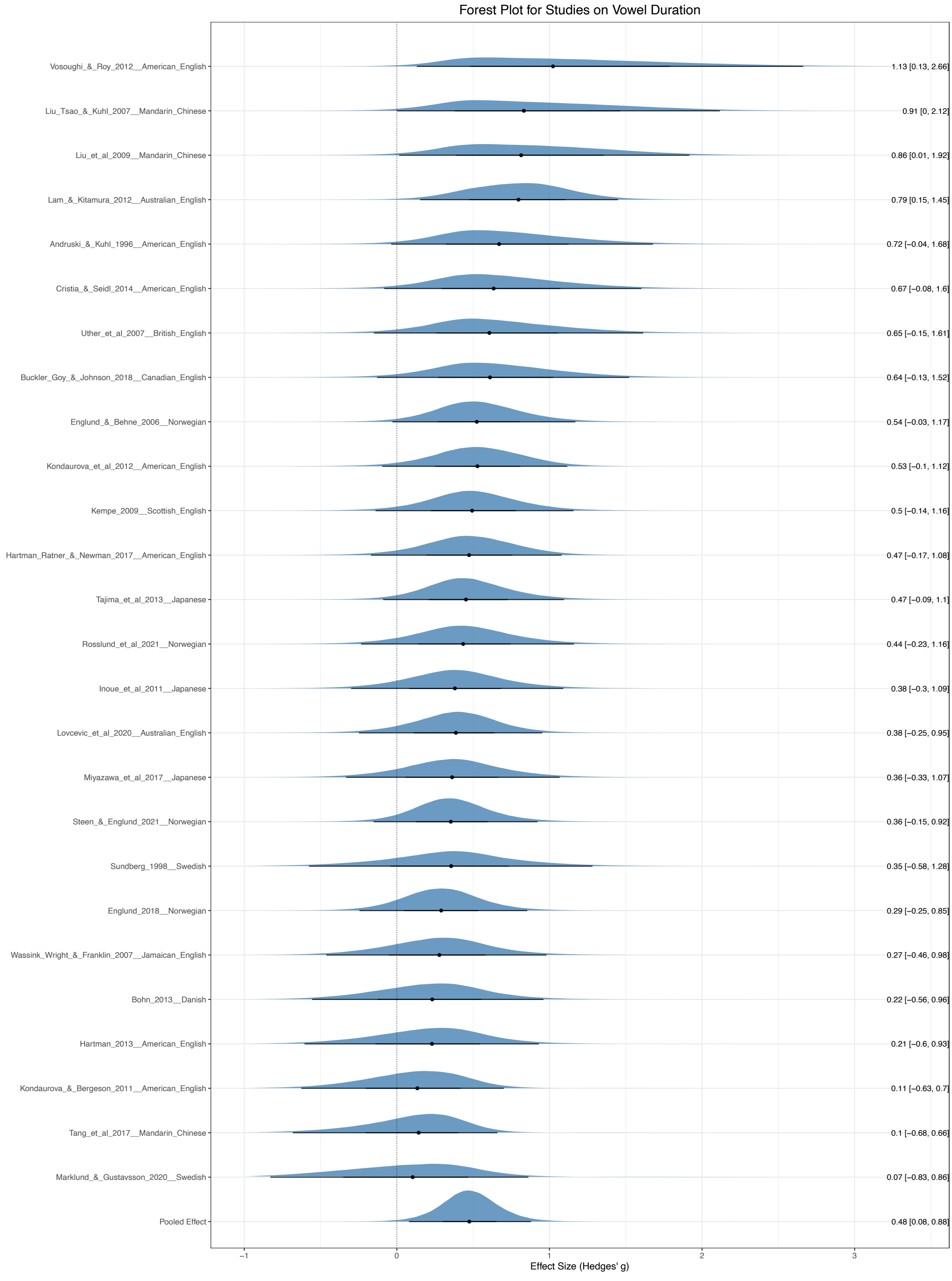
Supplementary Figure 6.2. Forest Plot for fo variability estimates according to study. The estimates are based on a total of 3006 participants across 44 studies investigating 34 distinct languages. The shaded areas indicate the posterior probability density of each estimate. The numbers to the right provide the estimated mean effect size (Hedges' *g*) and upper and lower 95% credible intervals).



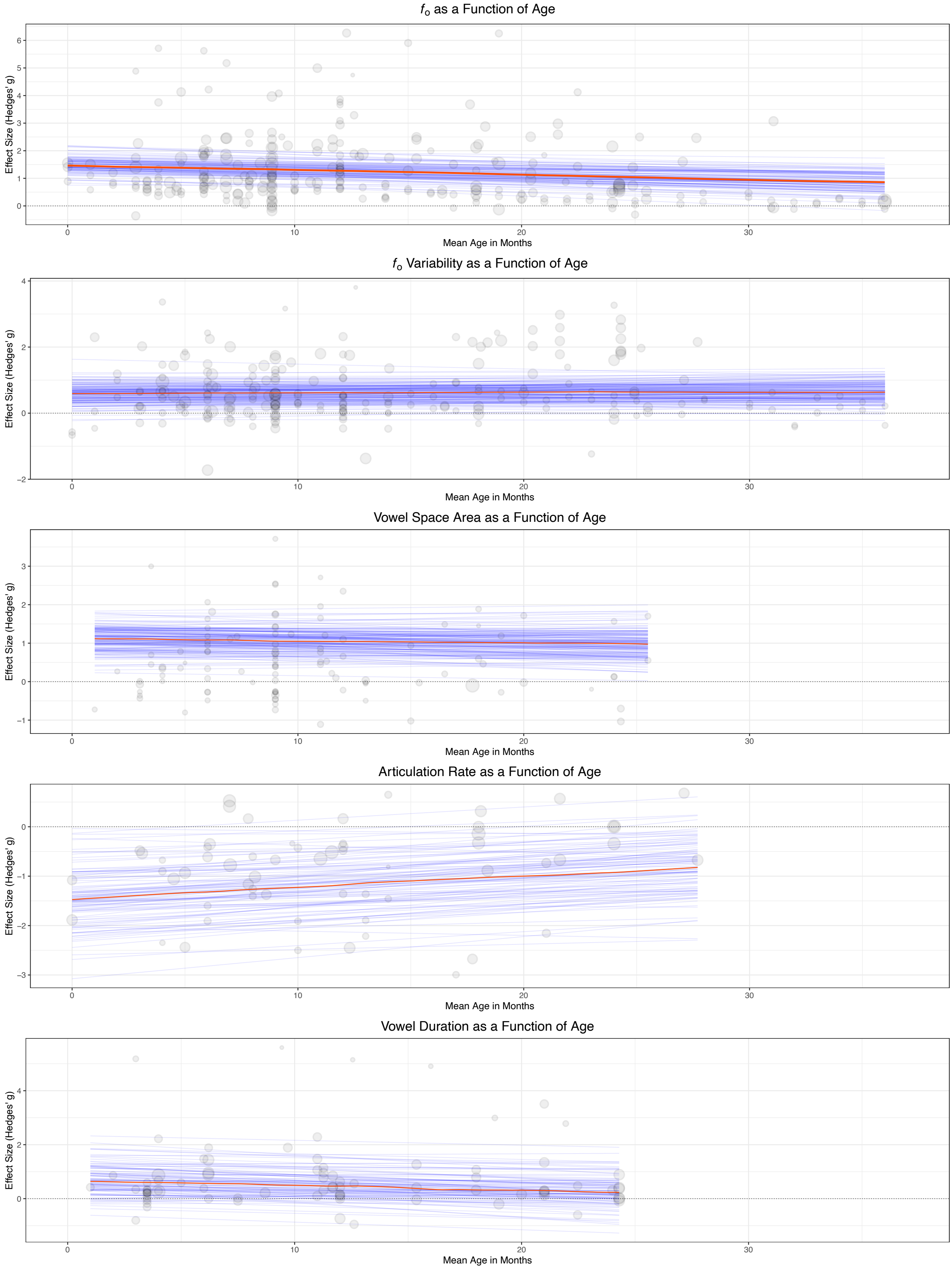
Supplementary Figure 6.3. Forest Plot for vowel space area estimates according to study. The estimates are based on a total of 1702 participants across 33 studies investigating 30 distinct languages. The shaded areas indicate the posterior probability density of each estimate. The numbers to the right provide the estimated mean effect size (Hedges' *g*) and upper and lower 95% credible intervals). The estimates within each study are broken down according to language; this is especially evident for studies[1, 54], from which data on a diverse set of languages exist.



Supplementary Figure 6.4. Forest plot for articulation rate estimates according to study. The estimates are based on a total of 976 participants across 17 studies investigating 17 distinct languages. The shaded areas indicate the posterior probability density of each estimate. The numbers to the right provide the estimated mean effect size (Hedges' *g*) and upper and lower 95% credible intervals).

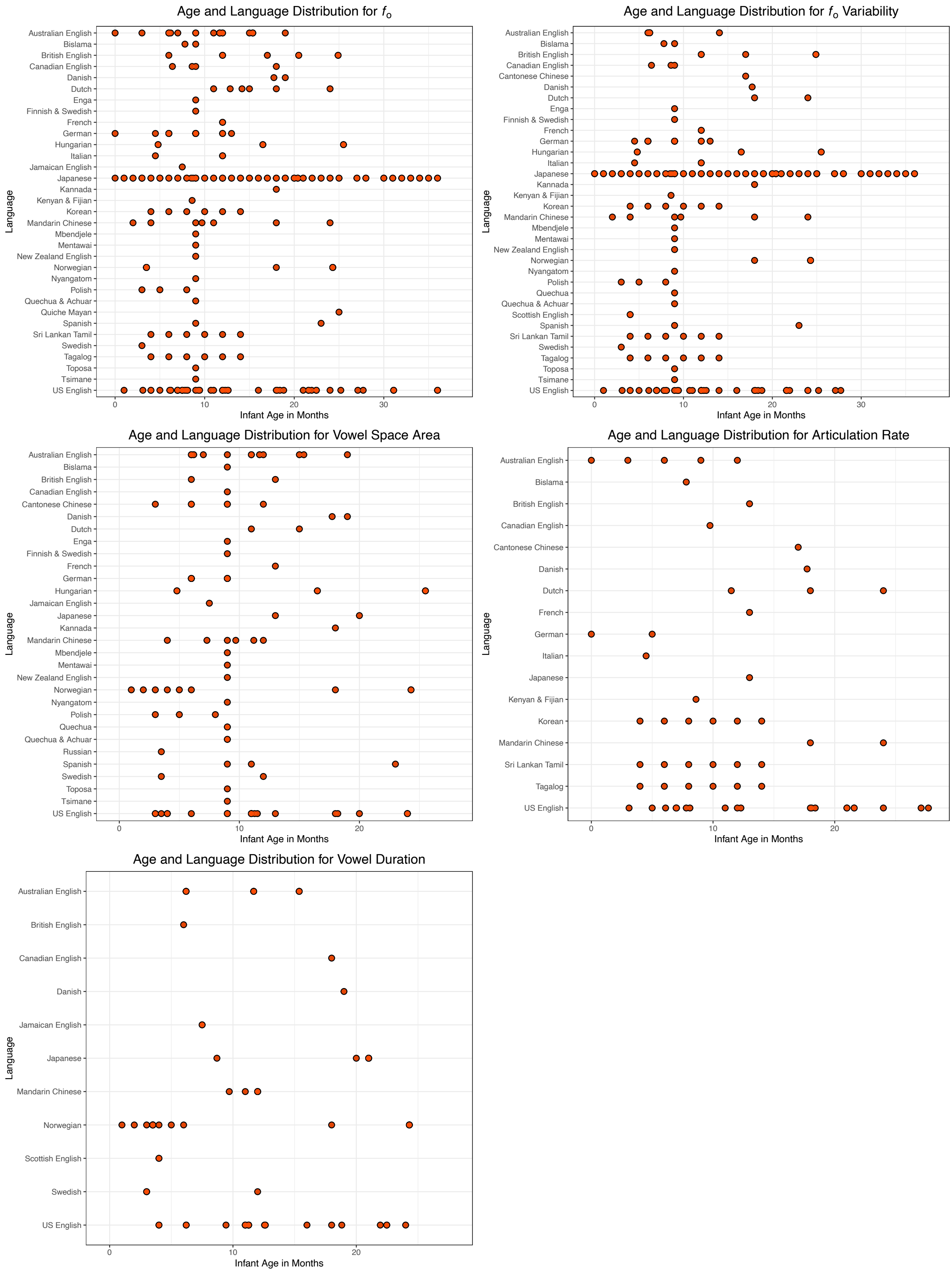


Supplementary Figure 6.5. Forest plot for vowel duration estimates according to study. The estimates are based on a total of 1411 participants across 26 studies investigating 11 distinct languages. The shaded areas indicate the posterior probability density of each estimate. The numbers to the right provide the estimated mean effect size (Hedges' *g*) and upper and lower 95% credible intervals).



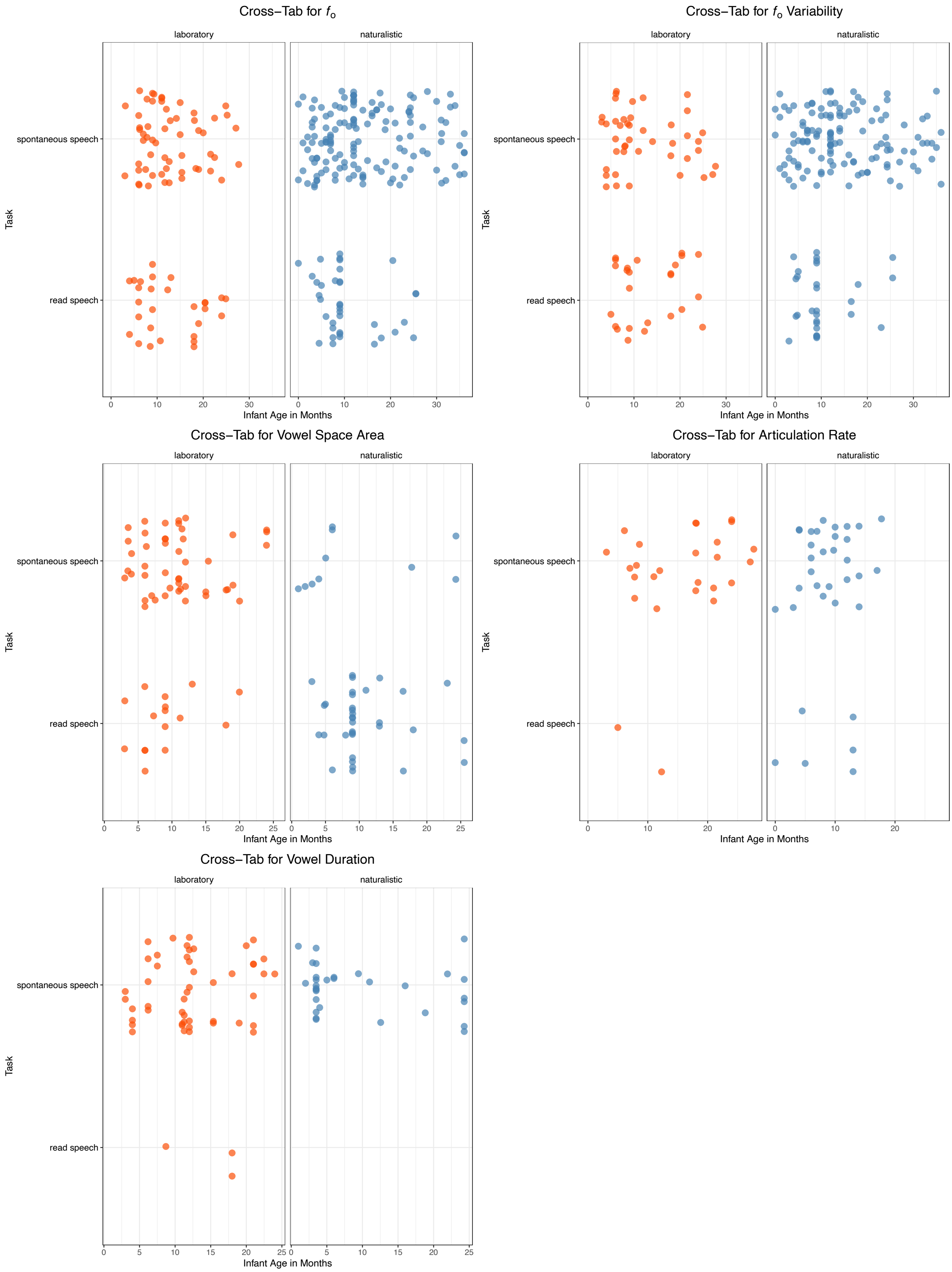
Supplementary Figure 6.6. A panel of plots showing each acoustic measure as a function of infant age. The blue lines reflect 100 posterior model predictions for the effect size estimates for each acoustic variable. As above, the grey points show the raw effect size measures. The size of the points is proportional to the inverse of the standard error of the effect size (i.e., the larger the point, the smaller the standard error). Note that each acoustic measure has an x-axis with different limits based on the data available.

S7: Age Distributions Across Languages



Supplementary Figure 7: Plots showing the age distribution by language of the effect sizes for each of the acoustic variables under investigation.

S8: Cross-Tab for Task & Environment



Supplementary Figure 8: Cross-tab for the predictors of task and environment with age on the x-axis. The colors are purely for aesthetic purposes: blue denotes recordings in a naturalistic environment, orange signifies recordings done in the lab.

S9: Parameter Estimates for Best Models

Model Parameters for f_0	**Estimate**
Standard Deviation of Languages	0.29 [0.01; 0.72]
Standard Deviation of Studies within Languages	0.91 [0.72; 1.14]
Standard Deviation of Measurements	0.07 [0.00; 0.19]
Age (months)	-0.02 [-0.03 0.01]
Australian English	1.29 [0.72; 1.84]
Bislama	0.85 [-0.37; 2.11]
British English	1.56 [0.43; 2.68]
Canadian English	1.21 [0.16; 2.21]
Danish	1.29 [-0.08; 2.64]
Dutch	0.97 [-0.12; 2.08]
Enga	1.33 [-0.25; 2.87]
Finnish & Swedish	1.19 [-0.27; 2.68]
French	1.53 [-0.01; 3.13]
German	1.34 [0.34; 2.35]
Hungarian	1.14 [-0.35; 2.62]
Italian	1.3 [0.03; 2.62]
Jamaican English	0.68 [-0.77; 2.18]
Japanese	1.04 [0.15; 1.99]
Kannada	1.46 [-0.17; 3.1]
Kenyan & Fijian	0.72 [-0.78; 2.22]
Korean	1 [-0.51; 2.47]

Mandarin Chinese	1.17 [0.15; 2.17]
Mbendjele	1.07 [-0.42; 2.57]
Mentawai	0.84 [-0.67; 2.35]
New Zealand English	1.65 [0.13; 3.09]
Norwegian	0.83 [-0.33; 1.99]
Nyangatom	1.37 [-0.15; 2.88]
Nu	9.41 [4.22; 21.65]

Supplementary Table 9.1: Model parameter estimates for f_0

Model parameters for f_0 variability	**Estimate**
Standard Deviation of Languages	0.21 [0.01; 0.58]
Standard Deviation of Studies within Languages	0.76 [0.60; 0.95]
Standard Deviation of Measurements	0.10 [0.01; 0.23]
Age (months)	0.00 [-0.01; 0.01]
Australian English	0.49 [0.03; 0.93]
Bislama	0.29 [-0.83; 1.4]
British English	0.57 [-0.45; 1.58]
Canadian English	0.82 [-0.16; 1.77]
Cantonese Chinese	0.51 [-0.93; 1.94]
Danish	1.21 [-0.19; 2.58]
Dutch	0.17 [-1.04; 1.45]
Enga	0.83 [-0.64; 2.3]

Finnish & Swedish	0.4 [-0.87; 1.73]
French	0.13 [-1.22; 1.5]
German	0.14 [-0.72; 1.01]
Hungarian	0.27 [-0.96; 1.51]
Italian	0.4 [-0.7; 1.48]
Japanese	0.55 [-0.16; 1.26]
Kannada	0.54 [-0.93; 2.04]
Kenyan & Fijian	0.55 [-0.81; 1.88]
Korean	0.31 [-0.99; 1.59]
Mandarin Chinese	0.42 [-0.47; 1.3]
Mbendjele	0.81 [-0.53; 2.13]
Mentawai	0.37 [-0.93; 1.67]
New Zealand English	0.55 [-0.7; 1.78]
Norwegian	0.89 [-0.16; 1.93]
Nyangatom	0.53 [-0.83; 1.85]
Polish	0.66 [-0.54; 1.85]
Quechua	-0.03 [-1.35; 1.29]
Quechua & Achuar	0.18 [-1.14; 1.49]
Scottish English	0.45 [-0.86; 1.71]
Spanish	-0.19 [-1.42; 1.06]
Sri Lankan Tamil	0.36 [-1.02; 1.69]

Swedish	0.16 [-1.3; 1.59]
Tagalog	0.26 [-1.02; 1.56]
Toposa	0.57 [-0.79; 1.93]
Tsimane	0.35 [-0.95; 1.67]
US English	0.92 [0.19; 1.56]
Spontaneous Speech	0.39 [0.05; 0.72]
Nu	31.17 [11.06; 67.91]

Supplementary Table 9.2: Model parameter estimates for f_0 variability

Model Parameters for Vowel Space Area	**Estimate**
Standard Deviation of Languages	0.37 [0.02; 0.87]
Standard Deviation of Studies within Languages	0.61 [0.41; 0.86]
Standard Deviation of Measurements	0.10 [0.00; 0.27]
Age (months)	-0.01 [-0.02; 0.01]
Australian English	0.93 [0.36; 1.42]
Bislama	1.94 [0.5; 3.29]
British English	0.99 [-0.32; 2.28]
Canadian English	1.54 [0.16; 2.83]
Cantonese Chinese	0.13 [-1.07; 1.41]
Danish	0.22 [-0.88; 1.36]
Dutch	-0.33 [-1.52; 0.96]

Enga	0.87 [-0.59; 2.32]
Finnish & Swedish	1.05 [-0.33; 2.37]
French	0.29 [-1.18; 1.79]
German	0.63 [-0.57; 1.86]
Hungarian	0.92 [-0.39; 2.23]
Jamaican English	0.53 [-0.76; 1.9]
Japanese	1.05 [-0.17; 2.23]
Kannada	1.24 [-0.33; 2.82]
Mandarin Chinese	0.84 [-0.14; 1.8]
Mbendjele	1.9 [0.36; 3.28]
Mentawai	0.21 [-1.14; 1.57]
New Zealand English	0.92 [-0.4; 2.2]
Norwegian	0.11 [-0.91; 1.18]
Nyangatom	0.96 [-0.44; 2.32]
Polish	0.41 [-0.82; 1.68]
Quechua	0.43 [-0.9; 1.79]
Quechua Achuar	1.82 [0.34; 3.17]
Russian	0.77 [-0.54; 2.12]
Spanish	0.5 [-0.7; 1.69]
Swedish	1.47 [0.28; 2.6]
Toposa	0.2 [-1.11; 1.56]
Tsimane	1.31 [-0.12; 2.67]
US English	0.66 [-0.24; 1.56]

Nu	18.61 [4.86; 50.08]

Supplementary Table 9.3: Model parameter estimates for vowel space area

Model parameters for articulation rate	**Estimate**
Standard Deviation of Languages	0.41 [0.02; 1.14]
Standard Deviation of Studies within Languages	0.74 [0.42; 1.19]
Standard Deviation of Measurements	0.23 [0.02; 0.46]
Age (months)	0.02 [0.00; 0.05]
Australian English	-1.19 [-2.14; -0.23]
Bislama	-0.49 [-2.06; 0.97]
British English	-1.53 [-3.32; 0.25]
Canadian English	-0.96 [-2.74; 0.92]
Cantonese Chinese	-1.97 [-3.77; -0.11]
Danish	-2.03 [-3.7; -0.25]
Dutch	-0.69 [-1.97; 0.49]
French	-1.33 [-3.08; 0.45]
German	-1.38 [-2.9; 0.14]
Italian	-1.18 [-2.81; 0.45]
Japanese	-1.65 [-3.42; 0.2]
Kenyan & Fijian	-1.26 [-2.74; 0.25]
Korean	-1.3 [-2.92; 0.32]

Mandarin Chinese	-0.56 [-2.03; 0.82]
Sri Lankan Tamil	-0.79 [-2.41; 0.82]
Tagalog	-1.5 [-3.11; 0.21]
US English	-0.9 [-2.03; 0.07]
Spontaneous Speech	0.95 [-0.08; 1.88]
Nu	23.64 [5.90; 58.56]

Supplementary Table 9.4: Model parameter estimates for articulation rate

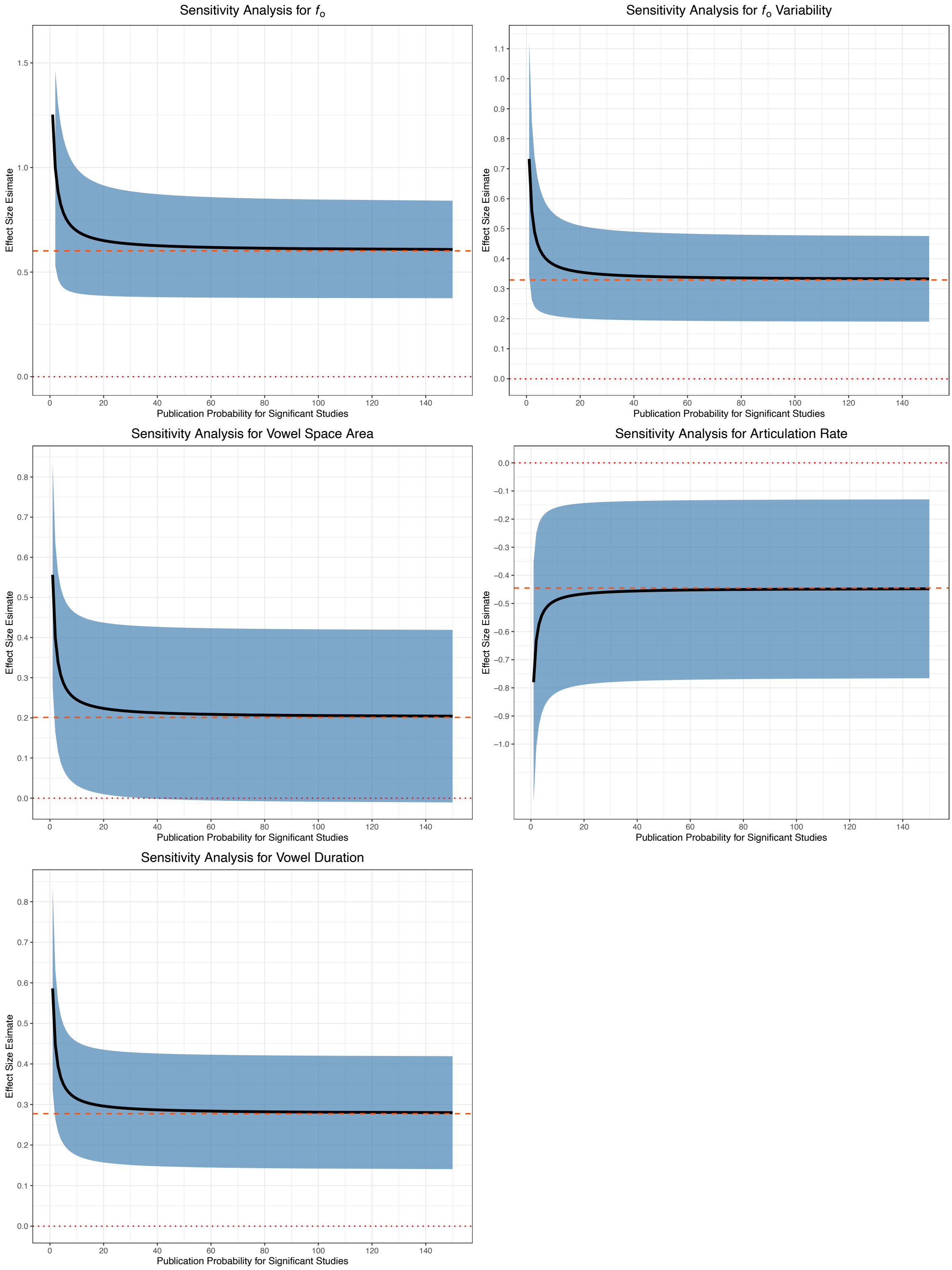
Model parameters for vowel duration	**Estimate**
Standard Deviation of Languages	0.39 [0.02; 1.14]
Standard Deviation of Studies within Languages	0.50 [0.12; 0.92]
Standard Deviation of Measurements	0.16 [0.01; 0.36]
Age (months)	-0.02 [-0.05; 0.01]
Australian English	0.6 [-0.08; 1.21]
British English	1.08 [-0.43; 2.42]
Canadian English	0.88 [-0.49; 2.18]
Danish	0.09 [-1.12; 1.44]
Jamaican English	0.17 [-1.04; 1.5]
Japanese	0.36 [-0.62; 1.36]
Mandarin Chinese	0.92 [-0.23; 1.98]
Norwegian	0.32 [-0.59; 1.35]
Scottish English	0.55 [-0.63; 1.7]
Swedish	-0.29 [-1.34; 1]

US English	0.84 [-0.18; 1.76]
Nu	4.90 [2.19; 11.19]

Supplementary Table 9.5: Model parameter estimates for vowel duration

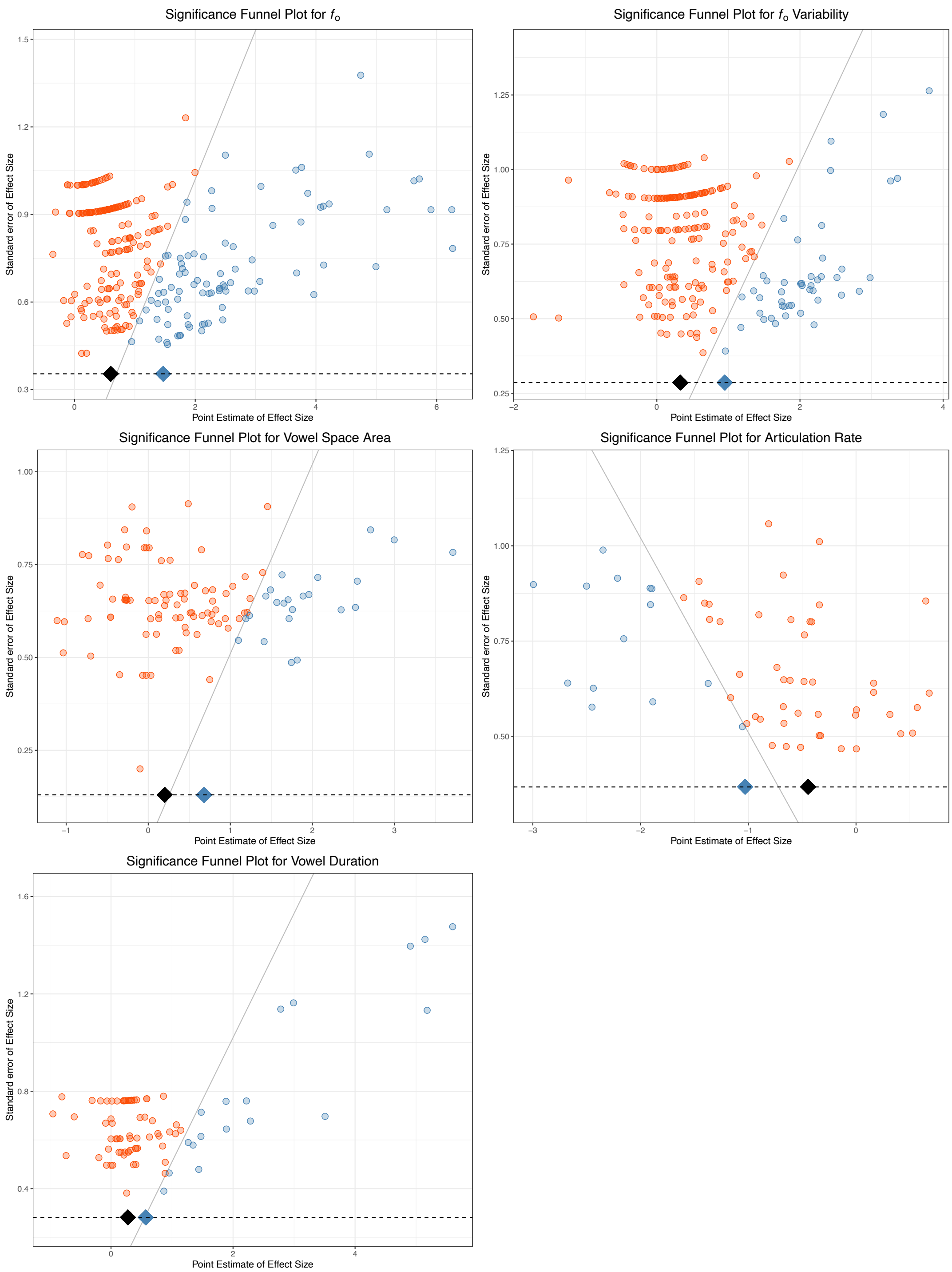
S10: Publication Bias Sensitivity Plots

The plot indicates what happens to the effect size if the publication probability is x times higher for significant studies than for non-significant studies. An effect size estimate of 0.0 is indicated by the orange dotted line, and the worst-case point estimate (see below) is indicated by the dashed red line.



Supplementary Figure 10.1: Sensitivity plots for each acoustic variable, showing the effect size estimate as a function of severity of publication bias.

Studies on the diagonal line have exactly p = 0.05. Black diamond: worst-case estimate of effect size based only on non-significant studies. Blue diamond: estimate of effect size for all studies. These plots help to determine the extent to which the non-affirmative studies' point estimates are systematically smaller than the entire set of point estimates. As a simple heuristic, when the diamonds are close to one another, our quantitative sensitivity analyses will typically indicate that the meta-analysis is fairly robust to publication bias. When the diamonds are distant or if the grey diamond represents a negligible effect size, then our sensitivity analyses may indicate that the meta-analysis is not robust.



Supplementary Figure 10.2: Significance funnel plots for each acoustic variable.

S11: Overview of Languages and Sample Sizes

Language	Total Sample Size
American English	2942
Australian English	1049
Bislama	36
British English	156
Canadian English	96
Cantonese Chinese	80
Danish	170
Dutch	335
French	30
German	710
Hungarian	234
Italian	110
Jamaican English	40
Japanese	1441
Kenyan & Fijian	45
Korean	87
Mandarin Chinese	373
Norwegian	924
Quiche Mayan	3
Russian	10
Scottish English	380
Spanish	17
Sri Lankan Tamil	84
Swedish	86
Tagalog	87

Supplementary Table 11.1 Overview of total sample size (i.e., number of speakers) according to language

S12: Overview of Studies, Measures & Number of Effect Sizes

Study	Measure	Language	Number of Effect Sizes
(Albin_&_Echols_1996)	F0	US English	4
(Amano_et_al_2006)	F0	Japanese	69
(Amano_et_al_2006)	F0V	Japanese	70
(Andruski_&_Kuhl_1996)	F0	US English	2
(Andruski_&_Kuhl_1996)	VD	US English	2
(Ashby_2004)	F0	Australian English	1
(Ashby_2004)	F0V	Australian English	1
(Benders_2013)	F0	Dutch	2
(Benders_2013)	VSA	Dutch	2
(Benders_StGeorge_&_Fletcher_2021)	F0	Dutch	2
(Bergeson_et_al_2006)	F0V	US English	2
(Bergeson_Miller_&_McCune_2006)	F0	US English	2
(Bohn_2013)	F0	Danish	1
(Bohn_2013)	VD	Danish	1
(Bohn_2013)	VSA	Danish	1
(Broen_1972)	AR	US English	2
(Broesch_&_Bryant_2015)	AR	Kenyan & Fijian	1
(Broesch_&_Bryant_2015)	AR	US English	1
(Broesch_&_Bryant_2015)	F0	Kenyan & Fijian	1
(Broesch_&_Bryant_2015)	F0	US English	1
(Broesch_&_Bryant_2015)	F0V	Kenyan & Fijian	1
(Broesch_&_Bryant_2015)	F0V	US English	1
(Broesch_&_Bryant_2018)	AR	Bislama	1
(Broesch_&_Bryant_2018)	AR	US English	1
(Broesch_&_Bryant_2018)	F0	Bislama	1
(Broesch_&_Bryant_2018)	F0	US English	1
(Broesch_&_Bryant_2018)	F0V	Bislama	1
(Broesch_&_Bryant_2018)	F0V	US English	1
(Buckler_Goy_&_Johnson_2018)	F0	Canadian English	1
(Buckler_Goy_&_Johnson_2018)	VD	Canadian English	1
(Burnham_et_al_2002)	F0	Australian English	1
(Burnham_et_al_2002)	VSA	Australian English	1
(Burnham_et_al_2015)	VSA	US English	7
(Church_2002)	AR	Canadian English	1
(Cooper_et_al_1997)	F0	US English	1
(Cooper_et_al_1997)	F0V	US English	1
(Cristia_&_Seidl_2014)	VD	US English	1

(Cristia_&_Seidl_2014)	VSA	US English	4
(Cristia_2010)	AR	US English	2
(Cristia_2010)	F0	US English	2
(Cristia_2010)	F0V	US English	2
(De_Palma_&_VanDam_2017)	F0	US English	2
(Dideriksen_&_Fusaroli_2018)	AR	Danish	1
(Dideriksen_&_Fusaroli_2018)	F0	Danish	1
(Dideriksen_&_Fusaroli_2018)	F0V	Danish	1
(Dideriksen_&_Fusaroli_2018)	VSA	Danish	1
(Dodane_&_Al-Tamimi_2007)	AR	British English	1
(Dodane_&_Al-Tamimi_2007)	AR	French	1
(Dodane_&_Al-Tamimi_2007)	AR	Japanese	1
(Dodane_&_Al-Tamimi_2007)	VSA	British English	1
(Dodane_&_Al-Tamimi_2007)	VSA	French	1
(Dodane_&_Al-Tamimi_2007)	VSA	Japanese	1
(Englund_&_Behne_2006)	VD	Norwegian	6
(Englund_&_Behne_2006)	VSA	Norwegian	6
(Englund_2018)	F0	Norwegian	6
(Englund_2018)	VD	Norwegian	12
(Fernald_&_Simon_1984)	AR	German	1
(Fernald_&_Simon_1984)	F0	German	1
(Fernald_1989)	F0	British English	2
(Fernald_1989)	F0	French	2
(Fernald_1989)	F0	German	2
(Fernald_1989)	F0	Italian	2
(Fernald_1989)	F0	Japanese	2
(Fernald_1989)	F0	US English	2
(Fernald_1989)	F0V	British English	2
(Fernald_1989)	F0V	French	2
(Fernald_1989)	F0V	German	2
(Fernald_1989)	F0V	Italian	2
(Fernald_1989)	F0V	Japanese	2
(Fernald_1989)	F0V	US English	2
(Fisher_&_Tokura_1995)	F0V	US English	1
(Garcia-Sierra_et_al_2021)	VSA	Spanish	1
(Gergely_et_al_2017)	F0	Hungarian	6
(Gergely_et_al_2017)	F0V	Hungarian	6
(Gergely_et_al_2017)	VSA	Hungarian	6
(Grieser_&_Kuhl_1988)	F0	Mandarin Chinese	1
(Grieser_&_Kuhl_1988)	F0V	Mandarin Chinese	1
(Han_De_Jong_&_Kager_2020)	F0	Dutch	2
(Han_De_Jong_&_Kager_2020)	F0	Mandarin Chinese	2
(Han_De_Jong_&_Kager_2020)	F0V	Dutch	2

(Han_De_Jong_&_Kager_2020)	F0V	Mandarin Chinese	2
(Han_et_al_2021)	AR	Dutch	2
(Han_et_al_2021)	AR	Mandarin Chinese	2
(Hartman_2013)	VD	US English	1
(Hartman_2013)	VSA	US English	2
(Hartman_Ratner_&_Newman_2017)	VD	US English	3
(Hartman_Ratner_&_Newman_2017)	VSA	US English	4
(Hilton_et_al_2022)	F0	Bislama	1
(Hilton_et_al_2022)	F0	Canadian English	1
(Hilton_et_al_2022)	F0	Enga	1
(Hilton_et_al_2022)	F0	Finnish & Swedish	1
(Hilton_et_al_2022)	F0	Kannada	1
(Hilton_et_al_2022)	F0	Mandarin Chinese	2
(Hilton_et_al_2022)	F0	Mbendjele	1
(Hilton_et_al_2022)	F0	Mentawai	1
(Hilton_et_al_2022)	F0	New Zealand English	1
(Hilton_et_al_2022)	F0	Nyangatom	1
(Hilton_et_al_2022)	F0	Polish	3
(Hilton_et_al_2022)	F0	Quechua & Achuar	1
(Hilton_et_al_2022)	F0	Spanish	2
(Hilton_et_al_2022)	F0	Toposa	1
(Hilton_et_al_2022)	F0	Tsimane	1
(Hilton_et_al_2022)	F0	US English	1
(Hilton_et_al_2022)	F0V	Bislama	1
(Hilton_et_al_2022)	F0V	Canadian English	1
(Hilton_et_al_2022)	F0V	Enga	1
(Hilton_et_al_2022)	F0V	Finnish & Swedish	1
(Hilton_et_al_2022)	F0V	Kannada	1
(Hilton_et_al_2022)	F0V	Mandarin Chinese	2
(Hilton_et_al_2022)	F0V	Mbendjele	1
(Hilton_et_al_2022)	F0V	Mentawai	1
(Hilton_et_al_2022)	F0V	New Zealand English	1
(Hilton_et_al_2022)	F0V	Nyangatom	1
(Hilton_et_al_2022)	F0V	Polish	3
(Hilton_et_al_2022)	F0V	Quechua	1
(Hilton_et_al_2022)	F0V	Quechua & Achuar	1
(Hilton_et_al_2022)	F0V	Spanish	2
(Hilton_et_al_2022)	F0V	Toposa	1
(Hilton_et_al_2022)	F0V	Tsimane	1
(Hilton_et_al_2022)	F0V	US English	1
(Hilton_et_al_2022)	VSA	Bislama	1
(Hilton_et_al_2022)	VSA	Canadian English	1
(Hilton_et_al_2022)	VSA	Enga	1

(Hilton_et_al_2022)	VSA	Finnish & Swedish	1
(Hilton_et_al_2022)	VSA	Kannada	1
(Hilton_et_al_2022)	VSA	Mandarin Chinese	2
(Hilton_et_al_2022)	VSA	Mbendjele	1
(Hilton_et_al_2022)	VSA	Mentawai	1
(Hilton_et_al_2022)	VSA	New Zealand English	1
(Hilton_et_al_2022)	VSA	Nyangatom	1
(Hilton_et_al_2022)	VSA	Polish	3
(Hilton_et_al_2022)	VSA	Quechua	2
(Hilton_et_al_2022)	VSA	Quechua & Achuar	2
(Hilton_et_al_2022)	VSA	Spanish	2
(Hilton_et_al_2022)	VSA	Toposa	1
(Hilton_et_al_2022)	VSA	Tsimane	1
(Hilton_et_al_2022)	VSA	US English	1
(Igarashi_et_al_2013)	F0	Japanese	3
(Igarashi_et_al_2013)	F0V	Japanese	3
(Ikeda_&_Masataka_1999)	F0	Japanese	1
(Ikeda_&_Masataka_1999)	F0V	Japanese	1
(Inoue_et_al_2011)	F0	Japanese	1
(Inoue_et_al_2011)	F0V	Japanese	1
(Inoue_et_al_2011)	VD	Japanese	1
(Kalashnikova_&_Burnham_2018)	F0	Australian English	5
(Kalashnikova_&_Burnham_2018)	VSA	Australian English	5
(Kalashnikova_et_al_2017)	VSA	Australian English	1
(Kalashnikova_et_al_2020)	F0	Australian English	1
(Kalashnikova_et_al_2020)	VSA	Australian English	1
(Kempe_2009)	F0V	Scottish English	2
(Kempe_2009)	VD	Scottish English	2
(Kondaurova_&_Bergeson_2011)	F0	US English	6
(Kondaurova_&_Bergeson_2011)	VD	US English	6
(Kondaurova_Bergeson_&_Dilley_2012)	VSA	US English	1
(Kondaurova_et_al_2012)	VD	US English	4
(Kondaurova_et_al_2013)	AR	US English	9
(Kondaurova_et_al_2013)	F0	US English	9
(Kondaurova_et_al_2013)	F0V	US English	11
(Kuhl_et_al_1997)	VSA	Russian	1
(Kuhl_et_al_1997)	VSA	Swedish	1
(Kuhl_et_al_1997)	VSA	US English	1
(Lahey_&_Ernestus_2014)	AR	Dutch	1
(Lam_&_Kitamura_2012)	F0	Australian English	1
(Lam_&_Kitamura_2012)	F0V	Australian English	3
(Lam_&_Kitamura_2012)	VD	Australian English	3
(Lam_&_Kitamura_2012)	VSA	Australian English	1

(Lee_et_al_2014)	AR	Australian English	5
(Lee_et_al_2014)	F0	Australian English	5
(Liu_et_al_2009)	F0V	Mandarin Chinese	1
(Liu_et_al_2009)	VD	Mandarin Chinese	1
(Liu_Kuhl_&_Tsao_2003)	VSA	Mandarin Chinese	2
(Liu_Tsao_&_Kuhl_2007)	F0	Mandarin Chinese	1
(Liu_Tsao_&_Kuhl_2007)	VD	Mandarin Chinese	1
(Liu_Tsao_&_Kuhl_2009)	F0	Mandarin Chinese	1
(Liu_Tsao_&_Kuhl_2009)	VSA	Mandarin Chinese	1
(Lovcevic_et_al_2020)	F0V	Australian English	1
(Lovcevic_et_al_2020)	VD	Australian English	6
(Lovcevic_et_al_2020)	VSA	Australian English	2
(Lovcevic_Kalashnikova_&_Burnham_2020)	F0	Australian English	6
(Marklund_&_Gustavsson_2020)	VD	Swedish	1
(Marklund_&_Gustavsson_2020)	VSA	Swedish	1
(Masataka_1992)	F0	Japanese	1
(Masataka_1992)	F0V	Japanese	1
(McMurray_et_al_2013)	F0	US English	1
(McMurray_et_al_2013)	F0V	US English	1
(Miyazawa_et_al_2017)	F0	Japanese	1
(Miyazawa_et_al_2017)	F0V	Japanese	1
(Miyazawa_et_al_2017)	VD	Japanese	1
(Miyazawa_et_al_2017)	VSA	Japanese	1
(Naoi_et_al_2012)	F0	Japanese	1
(Naoi_et_al_2012)	F0V	Japanese	1
(Narayan_&_McDermott_2016)	AR	Korean	6
(Narayan_&_McDermott_2016)	AR	Sri Lankan Tamil	6
(Narayan_&_McDermott_2016)	AR	Tagalog	6
(Narayan_&_McDermott_2016)	F0	Korean	6
(Narayan_&_McDermott_2016)	F0	Sri Lankan Tamil	6
(Narayan_&_McDermott_2016)	F0	Tagalog	6
(Narayan_&_McDermott_2016)	F0V	Korean	6
(Narayan_&_McDermott_2016)	F0V	Sri Lankan Tamil	6
(Narayan_&_McDermott_2016)	F0V	Tagalog	6
(Niwano_&_Sugai_2002)	F0	Japanese	4
(Niwano_&_Sugai_2002)	F0V	Japanese	4
(Ogle_&_Maidment_1993)	F0	British English	1
(Ogle_&_Maidment_1993)	F0V	British English	1
(Outters_et_al_2020)	F0V	German	2
(Outters_Schreiner_Behne_&_Mani_2020)	F0	German	2
(Raneri_2015)	AR	US English	4
(Raneri_2015)	F0	US English	4
(Raneri_2015)	F0V	US English	4

(Räsänen_et_al_2017)	F0	Canadian English	1
(Räsänen_et_al_2017)	F0V	Canadian English	1
(Ratner_&_Pye_1984)	F0	Quiche Mayan	1
(Ratner_&_Pye_1984)	F0	US English	1
(Rattanasone_Burnham_&_Reilly_2013)	VSA	Cantonese Chinese	6
(Rosslund_et_al_2021)	F0	Norwegian	1
(Rosslund_et_al_2021)	F0V	Norwegian	1
(Rosslund_et_al_2021)	VD	Norwegian	1
(Rosslund_et_al_2021)	VSA	Norwegian	1
(Sheehan_2008)	AR	US English	2
(Sheehan_2008)	F0	US English	4
(Sheehan_2008)	F0V	US English	4
(Shute_&_Wheldall_1989)	F0	British English	2
(Shute_&_Wheldall_2001)	F0	British English	2
(Shute_&_Wheldall_2001)	F0V	British English	2
(Smith_&_Strader_2014)	F0	US English	1
(Smith_&_Strader_2014)	F0V	US English	1
(Steen_&_Englund_2021)	F0	Norwegian	6
(Steen_&_Englund_2021)	F0V	Norwegian	6
(Steen_&_Englund_2021)	VD	Norwegian	6
(Steen_&_Englund_2021)	VSA	Norwegian	2
(Stern_et_al_1983)	F0V	US English	4
(Sulpizio_et_al_2018)	AR	German	1
(Sulpizio_et_al_2018)	AR	Italian	1
(Sulpizio_et_al_2018)	F0	German	1
(Sulpizio_et_al_2018)	F0	Italian	1
(Sulpizio_et_al_2018)	F0V	German	1
(Sulpizio_et_al_2018)	F0V	Italian	1
(Sundberg_1998)	F0	Swedish	1
(Sundberg_1998)	F0V	Swedish	1
(Sundberg_1998)	VD	Swedish	2
(Tajima_et_al_2013)	VD	Japanese	6
(Tang_&_Maidment_1996)	AR	Cantonese Chinese	1
(Tang_&_Maidment_1996)	F0V	Cantonese Chinese	1
(Tang_et_al_2017)	VD	Mandarin Chinese	6
(Tang_et_al_2017)	VSA	Mandarin Chinese	1
(Trainor_Austin_&_Desjardins_2000)	F0V	Canadian English	1
(Trainor_et_al_2000)	F0	Canadian English	1
(Uther_et_al_2007)	F0	British English	1
(Uther_et_al_2007)	VD	British English	1
(Uther_et_al_2007)	VSA	British English	1
(VanDam_&_De_Palma_2014)	F0	US English	2
(Vosoughi_&_Roy_2012)	F0	US English	5

(Vosoughi_&_Roy_2012)	F0V	US English	5
(Vosoughi_&_Roy_2012)	VD	US English	5
(Wassink_et_al_2007)	VSA	Jamaican English	1
(Wassink_Wright_&_Franklin_2007)	F0	Jamaican English	1
(Wassink_Wright_&_Franklin_2007)	VD	Jamaican English	2
(Weirich_&_Simpson_2019)	F0	German	12
(Weirich_&_Simpson_2019)	F0V	German	12
(Weirich_&_Simpson_2019)	VSA	German	12
(Wieland_et_al_2015)	VSA	US English	2
(Xu_et_al_2013)	VSA	Australian English	1

Supplementary Table 12.1 Overview of studies, measures, languages and number of effect sizes