

Meta-analytical estimates of the effect of voicing on vowel duration in English are biased (LabPhon17, Vancouver)

Stefano Coretta — Institute for Phonetics and Speech Processing (LMU Munich), s.coretta@lmu.de

08/07/2020

Contents

1	Overview	1
2	Read data	2
3	Get studies estimates	2
3.1	Studies that provide means and standard deviation	3
3.2	Studies that provide only means (with no standard deviation)	5
3.3	Studies with raw data	8
3.4	Estimated effects with error	10
4	Meta-analysis	10
4.1	Marginal posterior distributions	11
4.2	Shrinkage	13
5	Bias	16
6	Conclusions	18
	References	19

1 Overview

Across a wide variety of languages, vowels tend to be shorter when followed by voiceless consonants, and longer when followed by voiced ones. This phenomenon has been called the “voicing effect” (Mitleb 1982) or “pre-fortis clipping” (Wells 1990). Among the earliest traceable mentions to this phenomenon there are Meyer (1903) for English (cited in Lindblom 1967), Meyer (1904) for German, Meyer and Gombocz (1909) for Hungarian, and Gregoire (1911) for French (all cited in Maddieson and Gandour 1976). After these, a great number of studies further confirmed the existence of the effect in these languages and reported it in an ever increasing list of others.

English is the language that by far received the most attention in relation to the voicing effect (Heffner 1937; House and Fairbanks 1953; Lisker 1957, 1974; Zimmerman and Sapon 1958; Peterson and Lehiste 1960; House 1961; Sharf 1962, 1964; Lindblom 1967; Halle and Stevens 1967; Halle, Stevens, and Oppenheim 1967; Slis and Cohen 1969a, 1969b; Chen 1970; Klatt 1973; Raphael 1975; Umeda 1975; Javkin 1976; Port and Dalby 1982; Mack 1982; Luce and Charles-Luce 1985; Summers 1987; Kluender, Diehl, and Wright 1988; Jong 1991, 2004; Laeuffer 1992; Fowler 1992; Warren and Jacks 2005; Ko 2018; Glewwe 2018; Sanker 2019, among others).

In light of rising concerns for statistical power and publication bias (Roettger, Winter, and Baayen 2018; Kirby and Sonderegger 2018; Vasishth and Gelman 2019), it is fundamental to obtain a meta-analytical

estimate of the voicing effect based on previous work, and to investigate potential biases that can affect the estimate. A Bayesian meta-analysis was carried out on 17 Bayesian estimates of the effect of voicing on vowel duration (13 from word-final syllables, 4 from non-word-final syllables) obtained from a selection of 13 studies of English according to the method described in Nicenboim, Roettger, and Vasishth (2018).

Data and code can be found at <https://github.com/stefanocoretta/2020-labphon>.

2 Read data

```
stressed <- read_csv("./data/raw/stressed.csv") %>%
  filter(manner == "stop") %>%
  mutate(
    voice = factor(voice, levels = c("voiceless", "voiced")),
    n_syl = factor(n_syl, levels = c("mono", "di", "tri")),
    word_pos_2 = ifelse(word_pos == "medial", "medial", "final")
  ) %>%
  mutate_if(is.character, as.factor)

studies <- stressed %>%
  select(study, n_syl:word_pos_2) %>%
  unique() %>%
  # Hussein 1994 is not included in the analysis, Port 1981 tri is included with
  # Port 1981 di. Klatt 1973 is not included since they only provided grand means.
  filter(study != "hussein1994", study != "port1981_tri", study != "klatt1973_mono", study != "klatt1973_di")
  droplevels()

# ko18_data <- read_delim("./data/raw/ko2018.txt", delim = "\t")

iterations <- 4000

if (file.exists("./data/cache/estimated_ln_tbl.rds")) {
  estimated_ln_tbl <- readRDS("./data/cache/estimated_ln_tbl.rds")
}

if (file.exists("./data/cache/estimated_ln_sd_tbl.rds")) {
  estimated_ln_sd_tbl <- readRDS("./data/cache/estimated_ln_sd_tbl.rds")
}
```

3 Get studies estimates

The first step is to get the estimate (with respective error) of the voicing effect for each study individually. A log-normal distribution is used here as the distribution of the outcome variable (vowel duration), given evidence that the duration of speech intervals (segments, utterances, pauses) tend to be log-normally distributed (Rosen 2005; Ratnikova 2017; Gahl and Baayen 2019). For each study, the calculated estimate is the ratio between vowel duration when the vowel is followed by a voiceless consonant vs when the following consonant is voiced (the “voicing effect (VE) ratio”; for example, a VE ratio of 1.5 means that vowels are 1.5 times longer when the following consonant is voiced).

Depending on the nature of the information available from each study, the following methods are used to obtain the estimated VE ratio:

- **Studies that provide measures of central tendency (means) and dispersion (standard deviations):** Bayesian linear regressions with measurement error.
- **Studies that only provide measures of central tendency (means):** Bayesian linear regression.

- **Studies that provide raw data:** Bayesian linear regression fitted to the raw data.

3.1 Studies that provide means and standard deviation

Using measurement error models with studies that reported standard deviations together with the means. The model includes `mi()` terms, where the standard error is calculated as sd/\sqrt{N} (N is the number of observations from which the standard deviation was obtained). See <https://github.com/paul-buerkner/brms/issues/643#issuecomment-483339742>.

```
p81_di <- filter(stressed, study == "port1981_di") %>%
  mutate(obs = 1:n())
p81_mono <- filter(stressed, study == "port1981_mono") %>%
  mutate(obs = 1:n())

# Fitting a lognormal likelihood with (voice/pair) leads to many divergent transitions.

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sigma)
)

p81_di_ln_bm <- brm(
  v_duration | mi(sd/sqrt(n_obs)) ~
    voice,
  data = p81_di,
  prior = priors,
  family = lognormal,
  seed = 9899,
  file = "./data/cache/p81_di_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 20),
  iter = iterations,
  save_mevars = TRUE
)

p81_mono_ln_bm <- brm(
  v_duration | mi(sd/sqrt(n_obs)) ~
    voice,
  data = p81_mono,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/p81_mono_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 20),
  iter = iterations,
  save_mevars = TRUE
)

l85_medial <- filter(stressed, study == "luce1985_medial")
l85_final <- filter(stressed, study == "luce1985_final")

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
```

```

prior(normal(0, 0.1), class = sigma)
# prior(lkj(2), class = cor)
)

l85_medial_ln_bm <- brm(
  v_duration | mi(sd/sqrt(n_obs)) ~
    voice + (1|pair),
  data = l85_medial,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/l85_medial_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 20),
  iter = iterations,
  save_mevars = TRUE
)

l85_final_ln_bm <- brm(
  v_duration | mi(sd/sqrt(n_obs)) ~
    voice + (1|pair),
  data = l85_final,
  family = lognormal(),
  prior = priors,
  seed = 9899,
  file = "./data/cache/l85_final_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 20),
  iter = iterations,
  save_mevars = TRUE
)

d89 <- filter(stressed, study == "davis1989") %>%
  mutate(obs = 1:n())

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma),
  prior(lkj(2), class = cor)
)

d89_ln_bm <- brm(
  v_duration | mi(sd/sqrt(n_obs)) ~
    voice + (1+voice|pair) + (1+voice|speaker),
  data = d89,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/d89_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 20),
  iter = iterations
)

```

3.2 Studies that provide only means (with no standard deviation)

Most studies only provide means, so Bayesian regressions are fitted to the means only.

```
h37 <- filter(stressed, study == "heffner1937")

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)

h37_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = h37,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/h37_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)
```

```
hf53 <- filter(stressed, study == "housefairbanks1953")

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)

hf53_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = hf53,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/hf53_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)
```

```
z58 <- filter(stressed, study == "zimmerman1958")

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)
```

```

z58_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = z58,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/z58_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)

```

```

pl60 <- filter(stressed, study == "petersonlehiste1960")

```

```

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)

```

```

pl60_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = pl60,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/pl60_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)

```

```

s62_di <- filter(stressed, study == "sharf1962_di")
s62_mono <- filter(stressed, study == "sharf1962_mono")

```

```

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)

```

```

s62_di_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = s62_di,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/s62_di_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)

```

```
s62_mono_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = s62_mono,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/s62_mono_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)
```

```
c70 <- filter(stressed, study == "chen1970")
```

```
priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)
```

```
c70_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = c70,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/c70_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)
```

```
m82 <- filter(stressed, study == "mack1982")
```

```
priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)
```

```
m82_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = m82,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/m82_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)
```

```

192 <- filter(stressed, study == "laeuffer1992")

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicevoiced),
  prior(normal(0, 0.1), class = sd),
  prior(normal(0, 0.1), class = sigma)
)

192_ln_bm <- brm(
  v_duration ~
    voice + (1|pair),
  data = 192,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  file = "./data/cache/192_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)

```

3.3 Studies with raw data

```

# k18 <- filter(ko18_data, phoneme %in% c("AE1", "EH1", "IH1")) %>%
#   mutate(
#     voice = factor(voice, levels = c("-voice", "+voice")),
#     phoneme = factor(phoneme, levels = c("AE1", "EH1", "IH1"))
#   )
#
# contrasts(k18$phoneme) <- "contr.sum"

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = voicePvoice),
  prior(normal(0, 1), class = b, coef = phoneme1),
  prior(normal(0, 1), class = b, coef = phoneme2),
  prior(normal(0, 1), class = b, coef = voicePvoice:phoneme1),
  prior(normal(0, 1), class = b, coef = voicePvoice:phoneme2),
  prior(cauchy(0, 0.1), class = sd),
  prior(cauchy(0, 0.1), class = sigma),
  prior(lkj(2), class = cor)
)

k18_ln_bm <- brm(
  duration ~
    voice *
    phoneme +
    (1+voice|subj),
  data = k18,
  prior = priors,
  family = lognormal(),
  seed = 9899,
  control = list(adapt_delta = 0.999, max_treedepth = 20),
)

```



```

file = "./data/cache/k18_ln_bm"
)

contrasts(eng_durations$vowel) <- "contr.sum"
eng_durations_di <- filter(eng_durations, num_syl == "di")
eng_durations_mono <- filter(eng_durations, num_syl == "mono")

priors <- c(
  set_prior("normal(0, 3)", class = "Intercept"),
  set_prior("normal(0, 1)", class = "b", coef = "voicingvoiced"),
  set_prior("normal(0, 1)", class = "b", coef = "vowel1"),
  set_prior("normal(0, 1)", class = "b", coef = "vowel2"),
  set_prior("normal(0, 1)", class = "b", coef = "voicingvoiced:vowel1"),
  set_prior("normal(0, 1)", class = "b", coef = "voicingvoiced:vowel2"),
  set_prior("normal(0, 1)", class = "b", coef = "speech_rate_c"),
  set_prior("cauchy(0, 0.1)", class = "sd"),
  set_prior("lkj(2)", class = "cor"),
  set_prior("cauchy(0, 0.1)", class = "sigma")
)

c19_di_ln_bm <- brm(
  v1_duration ~
    voicing *
    vowel +
    speech_rate_c +
    (1 + voicing | speaker) +
    (1 | word),
  data = eng_durations_di,
  prior = priors,
  family = lognormal(),
  file = "./data/cache/c19_di_ln_bm",
  control = list(adapt_delta = .9999, max_treedepth = 15),
  iter = iterations
)

priors <- c(
  set_prior("normal(0, 3)", class = "Intercept"),
  set_prior("normal(0, 1)", class = "b", coef = "voicingvoiced"),
  set_prior("normal(0, 1)", class = "b", coef = "vowel1"),
  set_prior("normal(0, 1)", class = "b", coef = "vowel2"),
  set_prior("normal(0, 1)", class = "b", coef = "voicingvoiced:vowel1"),
  set_prior("normal(0, 1)", class = "b", coef = "voicingvoiced:vowel2"),
  set_prior("normal(0, 1)", class = "b", coef = "speech_rate_c"),
  set_prior("cauchy(0, 0.1)", class = "sd"),
  set_prior("lkj(2)", class = "cor"),
  set_prior("cauchy(0, 0.1)", class = "sigma")
)

c19_mono_ln_bm <- brm(
  v1_duration ~
    voicing *
    vowel +
    speech_rate_c +
    (1 + voicing | speaker) +

```

```

    (1 | word),
    family = lognormal(),
    data = eng_durations_mono,
    prior = priors,
    file = "./data/cache/c19_mono_ln_bm"
  )

```

3.4 Estimated effects with error

Now we make a data set with the estimated VE ratio and relative standard error from each study, to be used in the meta-analysis.

```

if (!file.exists("./data/cache/estimated_ln_tbl.rds")) {
  estimated_ln <- list()

  # These are in alphabetical order so that joining below does not complain
  # about different levels
  estimated_ln[["chen1970"]] <- fixef(c70_ln_bm)["voicevoiced",]
  estimated_ln[["coretta2019_di"]] <- fixef(c19_di_ln_bm)["voicingvoiced",]
  estimated_ln[["coretta2019_mono"]] <- fixef(c19_mono_ln_bm)["voicingvoiced",]
  estimated_ln[["davis1989"]] <- fixef(d89_ln_bm)["voicevoiced",]
  estimated_ln[["heffner1937"]] <- fixef(h37_ln_bm)["voicevoiced",]
  estimated_ln[["housefairbanks1953"]] <- fixef(hf53_ln_bm)["voicevoiced",]
  estimated_ln[["ko2018"]] <- fixef(k18_ln_bm)["voicePvoice",]
  estimated_ln[["laeufer1992"]] <- fixef(l92_ln_bm)["voicevoiced",]
  estimated_ln[["luce1985_final"]] <- fixef(l85_final_ln_bm)["voicevoiced",]
  estimated_ln[["luce1985_medial"]] <- fixef(l85_medial_ln_bm)["voicevoiced",]
  estimated_ln[["mack1982"]] <- fixef(m82_ln_bm)["voicevoiced",]
  estimated_ln[["petersonlehiste1960"]] <- fixef(pl60_ln_bm)["voicevoiced",]
  estimated_ln[["port1981_di"]] <- fixef(p81_di_ln_bm)["voicevoiced",]
  estimated_ln[["port1981_mono"]] <- fixef(p81_mono_ln_bm)["voicevoiced",]
  estimated_ln[["sharf1962_di"]] <- fixef(s62_di_ln_bm)["voicevoiced",]
  estimated_ln[["sharf1962_mono"]] <- fixef(s62_mono_ln_bm)["voicevoiced",]
  estimated_ln[["zimmerman1958"]] <- fixef(z58_ln_bm)["voicevoiced",]

  estimated_ln_tbl <- plyr::ldply(estimated_ln, .id = "study") %>%
    mutate(
      source = factor("original", levels = c("original", "meta-analysis")),
      study = factor(study)
    ) %>%
    left_join(y = studies)

  saveRDS(estimated_ln_tbl, file = "./data/cache/estimated_ln_tbl.rds")
}

```

4 Meta-analysis

We can now fit a meta-analytical model using a Bayesian linear regression with measurement error (added with `se(Est.Error)` in the left-hand side of the model formula).

```

priors <- c(
  prior(normal(0, 3), class = Intercept),
  prior(normal(0, 1), class = b, coef = syl_posnonMfinal),
  prior(normal(0, 0.1), class = sd)
)

```

```

)

meta_ln_bm <- brm(
  Estimate | se(`Est.Error`) ~ syl_pos + (1 | study),
  data = estimated_ln_tbl,
  prior = priors,
  control = list(adapt_delta = 0.9999, max_treedepth = 15),
  seed = 9899,
  file = "./data/cache/meta_ln_bm"
)

summary(meta_ln_bm)

## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: Estimate | se(Est.Error) ~ syl_pos + (1 | study)
## Data: estimated_ln_tbl (Number of observations: 17)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Group-Level Effects:
## ~study (Number of levels: 17)
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)    0.12     0.03    0.08    0.18 1.00    1126    1690
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept         0.40     0.04    0.33    0.48 1.00     892    1438
## syl_posnonMfinal   -0.25     0.08   -0.40   -0.09 1.00    1379    1999
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).

meta_ln_bm_est <- fixef(meta_ln_bm)["Intercept", "Estimate"]
meta_ln_bm_q2.5 <- fixef(meta_ln_bm)["Intercept", "Q2.5"]
meta_ln_bm_q97.5 <- fixef(meta_ln_bm)["Intercept", "Q97.5"]

post_ln_nonfi <- (c(posterior_samples(meta_ln_bm, pars = "b_Intercept")) +
  posterior_samples(meta_ln_bm, pars = "b_syl_posnonMfinal"))

meta_ln_bm_syl_q2.5 <- quantile(post_ln_nonfi$b_syl_posnonMfinal, c(0.025))
meta_ln_bm_syl_q97.5 <- quantile(post_ln_nonfi$b_syl_posnonMfinal, c(0.975))
meta_ln_bm_syl_q50 <- (meta_ln_bm_syl_q2.5 + meta_ln_bm_syl_q97.5) / 2

```

4.1 Marginal posterior distributions

```

marginal_ln_post <- posterior_samples(meta_ln_bm, pars = "b_") %>%
  mutate(
    final = exp(b_Intercept),
    `non-final` = exp(b_Intercept + b_syl_posnonMfinal)
  ) %>%
  select(final:`non-final`) %>%

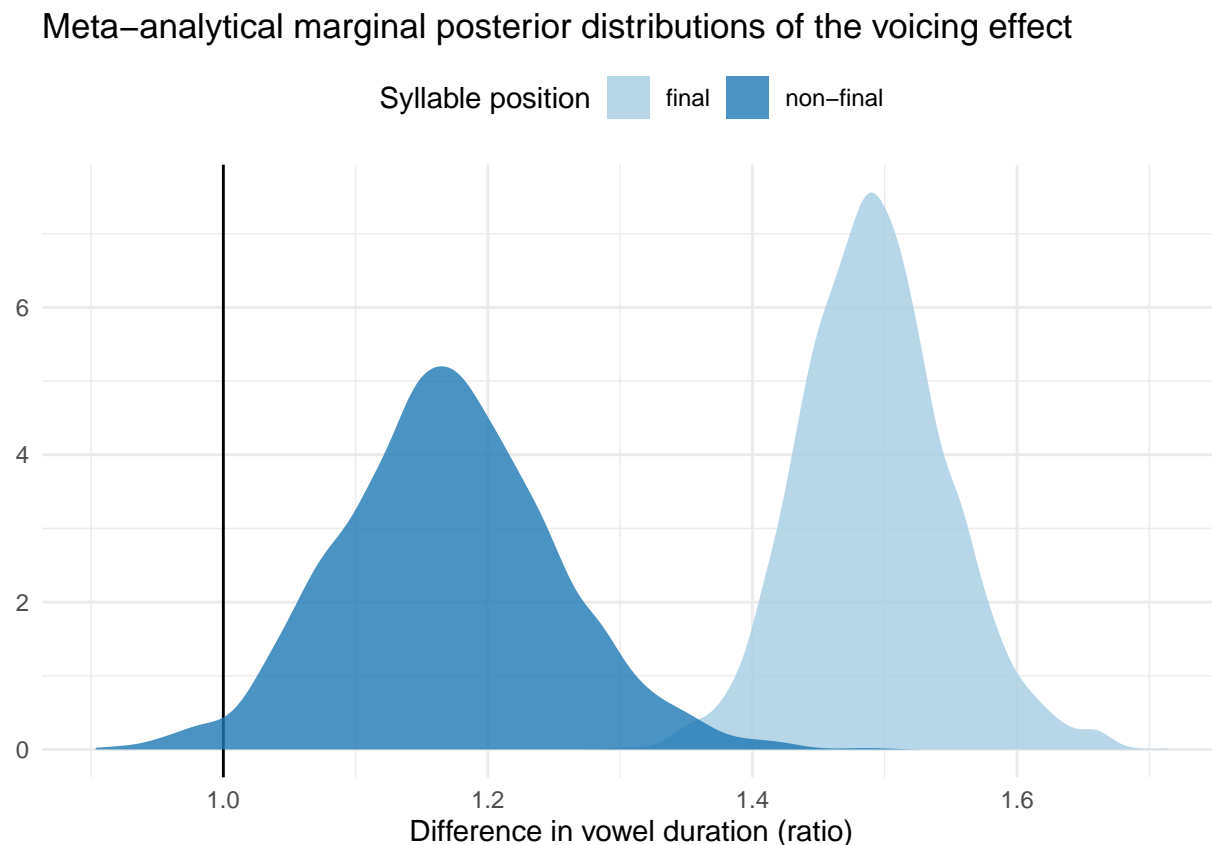
```

```

pivot_longer(everything(), names_to = "position", values_to = "posterior")

marginal_ln_post %>%
  ggplot(aes(posterior, fill = position)) +
  geom_vline(aes(xintercept = 1)) +
  geom_density(colour = NA, alpha = 0.8) +
  scale_fill_brewer(
    type = "qual", palette = "Paired"
  ) +
  labs(
    title = "Meta-analytical marginal posterior distributions of the voicing effect",
    x = "Difference in vowel duration (ratio)",
    fill = "Syllable position",
    y = element_blank()
  ) +
  theme(legend.position = "top")

```



```

marginal_ln_post %>%
  group_by(position) %>%
  summarise(
    lower.CI = quantile(posterior, p = 0.025),
    upper.CI = quantile(posterior, p = 0.975)
  )

```

```

## # A tibble: 2 x 3
##   position lower.CI upper.CI
##   <chr>      <dbl>   <dbl>

```

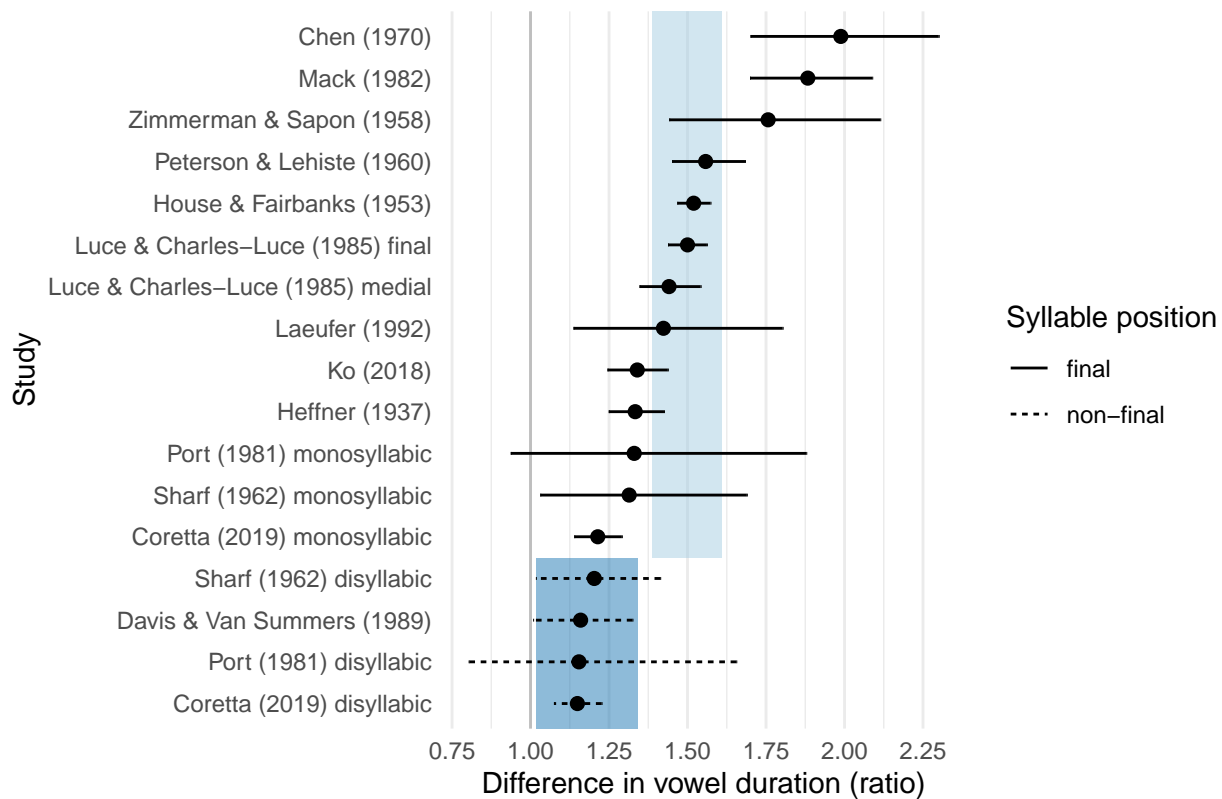
```
## 1 final          1.39      1.61
## 2 non-final      1.02      1.34
```

The 95% CI of the meta-analytical posterior of the voicing effect is [1.4, 1.6] in word-final position, and [1, 1.35] in penultimate position.

4.2 Shrinkage

```
estimated_ln_tbl %>%
  ggplot(aes(reorder(study_ref, Estimate), exp(Estimate), linetype = syl_pos)) +
  geom_hline(aes(yintercept = 1), colour = "grey") +
  scale_x_discrete() +
  annotate("rect", xmin = 4.5, xmax = Inf, ymin = exp(meta_ln_bm_q2.5), ymax = exp(meta_ln_bm_q97.5), a
  annotate("rect", xmin = -Inf, xmax = 4.5, ymin = exp(meta_ln_bm_syl_q2.5), ymax = exp(meta_ln_bm_syl_q
  geom_point(size = 2, position = position_dodge(width = 0.8)) +
  geom_errorbar(aes(ymin = exp(Q2.5), ymax = exp(Q97.5)), width = 0, position = position_dodge(width = 0
  scale_color_brewer(type = "qual", palette = "Dark2", breaks = c("original", "meta-analysis")) +
  scale_y_continuous(breaks = seq(0.5, 2.5, by = 0.25)) +
  scale_linetype_discrete(labels = c("final", "non-final")) +
  labs(
    # caption = "The shaded areas indicate the 95% CI of the meta-analytical posterior\nof the voicing
    title = "Studies' estimates of the Voicing Effect",
    y = "Difference in vowel duration (ratio)",
    x = "Study",
    linetype = "Syllable position",
    colour = "Source"
  ) +
  coord_flip() +
  theme(panel.grid.major.y = element_blank(), panel.grid.minor.y = element_blank())
```

Studies' estimates of the Voicing Effect



```
ggsave("./img/origin-3.png", width = 7, height = 4.5)
```

```
# Code adapted from Nicenboim et al. 2018
```

```
# We need this since the model includes `num_syl`
```

```
post_nonf <- posterior_samples(meta_ln_bm, pars = "b_syl_posnonMfinal")
```

```
studies_shrunk <- (c(posterior_samples(meta_ln_bm, pars = "b_Intercept")) +  
  posterior_samples(meta_ln_bm, pars = "r_study")) %>%  
  bind_cols(post_nonf) %>%
```

```
# The following mutate sums the random coefficients of the studies with  
# `num_syl` == "non-final" to the posterior samples of the non_syl term
```

```
mutate(  
  `r_study[davis1989,Intercept]` = `r_study[davis1989,Intercept]` + b_syl_posnonMfinal,  
  `r_study[sharf1962_di,Intercept]` = `r_study[sharf1962_di,Intercept]` + b_syl_posnonMfinal,  
  `r_study[coretta2019_di,Intercept]` = `r_study[coretta2019_di,Intercept]` + b_syl_posnonMfinal,  
  `r_study[port1981_di,Intercept]` = `r_study[port1981_di,Intercept]` + b_syl_posnonMfinal,  
) %>%
```

```
select(-b_syl_posnonMfinal) %>%
```

```
summarise_all(list(~list(c(  
  mean(.),  
  quantile(., probs = c(.025, 0.975)),  
  sd(.))
```

```
)))) %>%
```

```
unnest(cols = c(`r_study[chen1970,Intercept]`, `r_study[coretta2019_di,Intercept]`,  
  `r_study[coretta2019_mono,Intercept]`, `r_study[davis1989,Intercept]`,  
  `r_study[heffner1937,Intercept]`, `r_study[housefairbanks1953,Intercept]`,
```

```

`r_study[ko2018,Intercept]`, `r_study[laeufer1992,Intercept]`,
`r_study[luce1985_final,Intercept]`, `r_study[luce1985_medial,Intercept]`,
`r_study[mack1982,Intercept]`, `r_study[petersonlehiste1960,Intercept]`,
`r_study[port1981_di,Intercept]`, `r_study[port1981_mono,Intercept]`,
`r_study[sharf1962_di,Intercept]`, `r_study[sharf1962_mono,Intercept]`,
`r_study[zimmerman1958,Intercept]`)) %>%
transpose() %>%
setNames(c("Estimate", "Q2.5", "Q97.5", "Est.Error")) %>%
map_df(unlist) %>%
mutate(
  study = estimated_ln_tbl$study,
  source = factor("meta-analysis", levels = c("original", "meta-analysis"))
) %>%
left_join(y = studies)

```

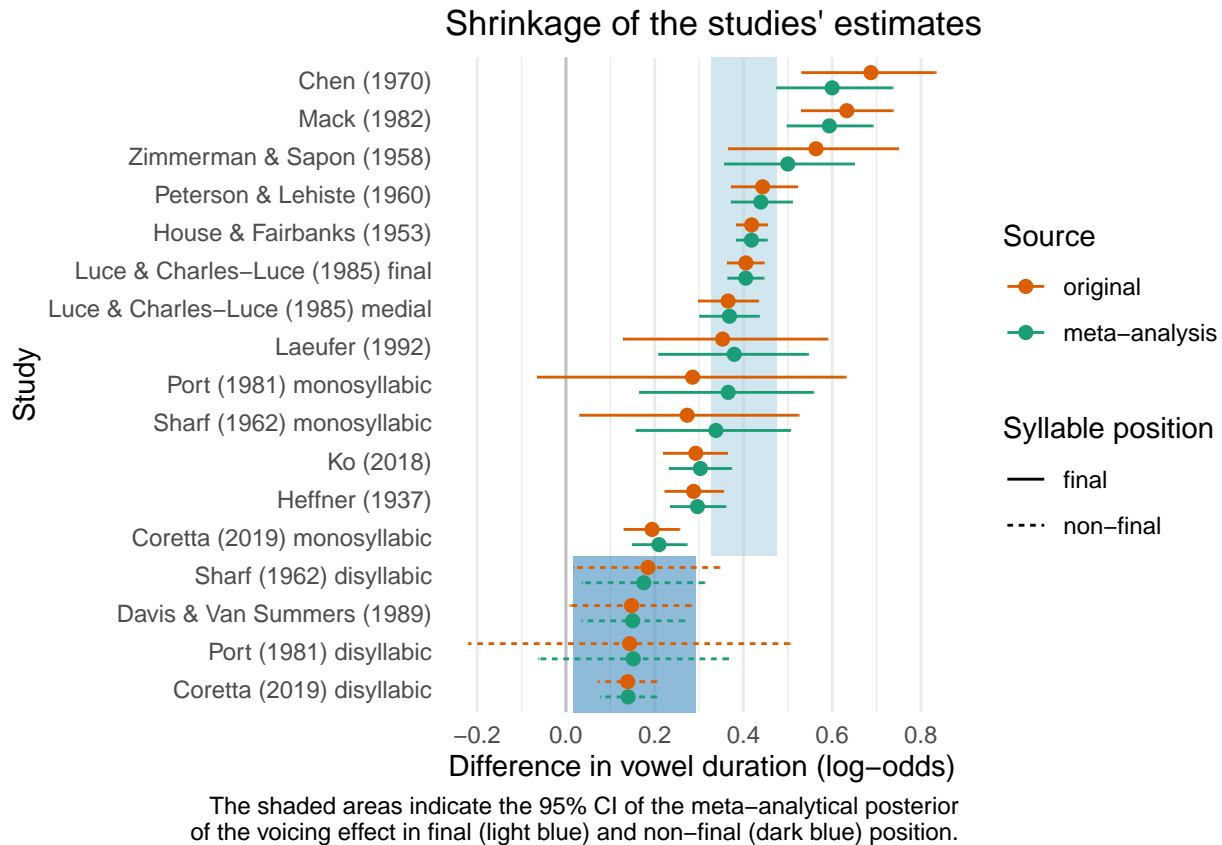
```
## Joining, by = "study"
```

The following plot shows the level of shrinkage of the individual studies' estimates in the meta-analysis. Estimates with greater uncertainty (greater error) shrunk more towards the meta-analytical estimate.

```

bind_rows(estimated_ln_tbl, studies_shrunk) %>%
mutate(source = factor(source, levels = c("meta-analysis", "original"))) %>%
ggplot(aes(reorder(study_ref, Estimate), Estimate, colour = source, linetype = syl_pos)) +
geom_hline(aes(yintercept = 0), colour = "grey") +
scale_x_discrete() +
annotate("rect", xmin = 4.5, xmax = Inf, ymin = meta_ln_bm_q2.5, ymax = meta_ln_bm_q97.5, alpha = 0.5) +
annotate("rect", xmin = -Inf, xmax = 4.5, ymin = meta_ln_bm_syl_q2.5, ymax = meta_ln_bm_syl_q97.5, alpha = 0.5) +
geom_point(size = 2, position = position_dodge(width = 0.8)) +
geom_errorbar(aes(ymin = Q2.5, ymax = Q97.5), width = 0, position = position_dodge(width = 0.8)) +
scale_color_brewer(type = "qual", palette = "Dark2", breaks = c("original", "meta-analysis")) +
scale_y_continuous(breaks = c(-0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1)) +
scale_linetype_discrete(labels = c("final", "non-final")) +
labs(
  title = "Shrinkage of the studies' estimates",
  caption = "The shaded areas indicate the 95% CI of the meta-analytical posterior\nof the voicing effect",
  y = "Difference in vowel duration (log-odds)",
  x = "Study",
  linetype = "Syllable position",
  colour = "Source"
) +
coord_flip() +
theme(panel.grid.major.y = element_blank(), panel.grid.minor.y = element_blank())

```



5 Bias

Potential biases in the meta-analytical estimate of the VE ratio are assessed with the funnel plot method (Light and Pillemer 1984). For a review of the method see Sterne, Becker, and Egger (2005) and Egger et al. (1997), for a critique Lau et al. (2006).

The following funnel plot shows the studies's mean estimates on the x -axis and their estimated error on the y -axis. The shaded light blue area is the 95% CI of the meta-analytical estimate of the voicing effect in word-final position, while the vertical line is the meta-analytical mean. In presence of bias, the dots in the funnel plot are not symmetrically distributed around the meta-analytical mean. (The number of studies for the word-non-final position are not sufficient to assess potential bias).

```
e.max <- by(estimated_ln_tbl$Est.Error, estimated_ln_tbl$syl_pos, range)[[1]][2]
se_range <- seq(0, e.max, by = 0.001)
ci <- tibble(
  x_seq = se_range,
  ci_lo = meta_ln_bm_est - 1.96 * se_range,
  ci_up = meta_ln_bm_est + 1.96 * se_range
)

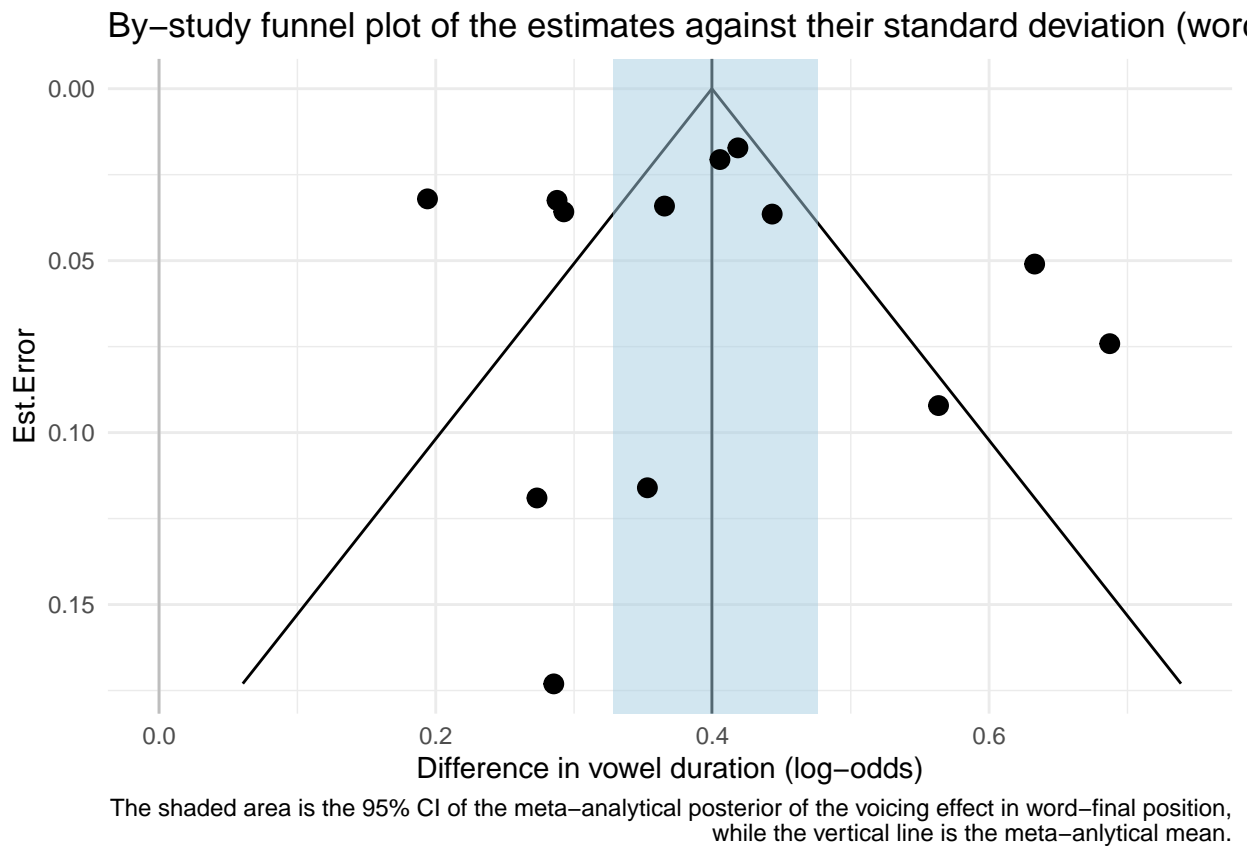
estimated_ln_tbl %>%
  filter(syl_pos == "final") %>%
  ggplot(aes(Estimate, Est.Error)) +
  geom_line(aes(y = x_seq, x = ci_up), data = ci) +
  geom_line(aes(y = x_seq, x = ci_lo), data = ci) +
  geom_vline(aes(xintercept = 0), colour = "grey") +
  geom_vline(aes(xintercept = meta_ln_bm_est)) +
```



```

annotate("rect", ymin = -Inf, ymax = Inf, xmin = meta_ln_bm_q2.5, xmax = meta_ln_bm_q97.5, alpha = 0.1) +
geom_point(size = 3) +
labs(
  title = "By-study funnel plot of the estimates against their standard deviation (word-final position)",
  caption = "The shaded area is the 95% CI of the meta-analytical posterior of the voicing effect in word-final position",
  x = "Difference in vowel duration (log-odds)"
) +
scale_y_reverse()

```



```

ggsave("./img/funnel-sd.png", width = 7, height = 4.5)

```

There is some indication of bias towards larger values (there are two studies with x higher than 0.6).

Another way is to use precision instead of standard deviation.

```

e.min <- by(estimated_ln_tbl$Est.Error, estimated_ln_tbl$syl_pos, range)[[1]][1]
e.max <- by(estimated_ln_tbl$Est.Error, estimated_ln_tbl$syl_pos, range)[[1]][2]
se_range <- seq(e.min, e.max, by = 0.001)
ci <- tibble(
  y_seq = 1/(seq(e.min, e.max, by = 0.001)^2),
  ci_lo = meta_ln_bm_est - 1.96 * se_range,
  ci_up = meta_ln_bm_est + 1.96 * se_range
)

# plot(seq(0, 30, by = 1) * 1.96, 1/seq(0, 30, by = 1))

estimated_ln_tbl %>%
  mutate(precision = 1/(Est.Error^2)) %>%

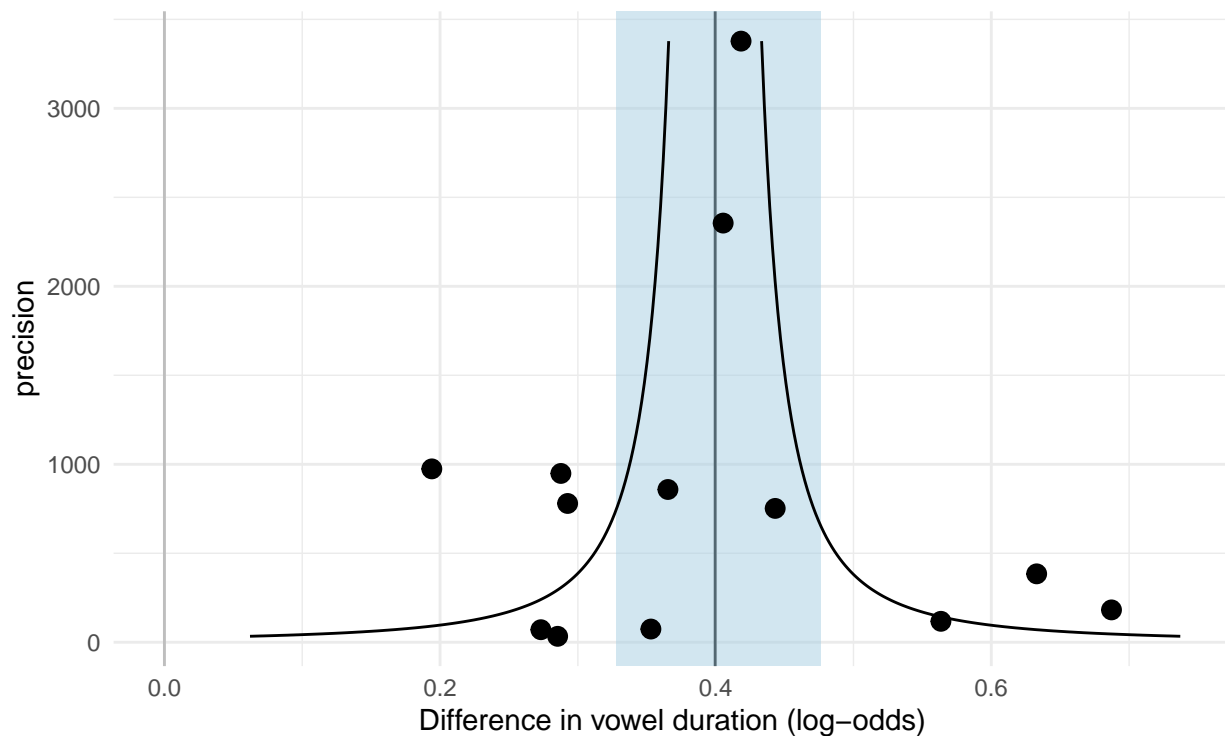
```

```

filter(syl_pos == "final") %>%
ggplot(aes(Estimate, precision)) +
geom_vline(aes(xintercept = 0), colour = "grey") +
geom_vline(aes(xintercept = meta_ln_bm_est)) +
annotate("rect", ymin = -Inf, ymax = Inf, xmin = meta_ln_bm_q2.5, xmax = meta_ln_bm_q97.5, alpha = 0.1) +
geom_point(size = 3) +
scale_x_continuous(breaks = seq(-1, 1, 0.2)) +
labs(
  title = "By-study funnel plot of the estimate against their precision (word-final position)",
  caption = "The shaded area is the 95% CI of the meta-analytical posterior of the voicing effect in word-final position",
  x = "Difference in vowel duration (log-odds)"
) +
geom_line(aes(y = y_seq, x = ci_up), data = ci) +
geom_line(aes(y = y_seq, x = ci_lo), data = ci)

```

By-study funnel plot of the estimate against their precision (word-final position)



The shaded area is the 95% CI of the meta-analytical posterior of the voicing effect in word-final position, while the vertical line is the meta-analytical mean.

6 Conclusions

A meta-analytical study of the voicing effect in English, based on 13 estimates, suggests possible publication bias (with estimates skewed towards greater values). The low number of publications that provide numeric summaries of the effect makes it difficult to draw more certain conclusions. Moreover, most studies only reported means without related measures of dispersions, which unduly increases the overall certainty of the estimates. More work on the voicing effect in English is needed to obtain a more certain estimate of the effect. Sharing data and analysis code in future studies is also paramount to aid future meta-analytical estimates.

References

- Chen, Matthew. 1970. "Vowel Length Variation as a Function of the Voicing of the Consonant Environment." *Phonetica* 22 (3): 129–59.
- Egger, Matthias, George Davey Smith, Martin Schneider, and Christoph Minder. 1997. "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *British Medical Journal* 315: 629–34.
- Fowler, Carol A. 1992. "Vowel Duration and Closure Duration in Voiced and Unvoiced Stops: There Are No Contrast Effects Here." *Journal of Phonetics* 20 (1): 143–65.
- Gahl, Susanne, and R. Harald Baayen. 2019. "Twenty-Eight Years of Vowels: Tracking Phonetic Variation Through Young to Middle Age Adulthood." *Journal of Phonetics* 74: 42–54.
- Glewwe, Eleanor. 2018. "The Effect of Lexical Competition on Vowel Duration Before Voiced and Voiceless English Stops." *The Journal of the Acoustical Society of America* 143 (3): 1968–8.
- Gregoire, A. 1911. "Influences Des Consonnes Occlusives Sur La Durée Des Syllabes Precedents." *Revue de Phonétique* 1: 260–92.
- Halle, Morris, and Kenneth Noble Stevens. 1967. "Mechanism of Glottal Vibration for Vowels and Consonants." *The Journal of the Acoustical Society of America* 41 (6): 1613–3.
- Halle, Morris, Kenneth Noble Stevens, and Alan Victor Oppenheim. 1967. "On the Mechanism of Glottal Vibration for Vowels and Consonants." In *Quarterly Progress Report*, 85:267–77.
- Heffner, R.-M.S. 1937. "Notes on the Length of Vowels." *American Speech* 12: 128–34.
- House, Arthur S. 1961. "On Vowel Duration in English." *The Journal of the Acoustical Society of America* 33 (9): 1174–8.
- House, Arthur S., and Grant Fairbanks. 1953. "The Influence of Consonant Environment Upon the Secondary Acoustical Characteristics of Vowels." *The Journal of the Acoustical Society of America* 25 (1): 105–13.
- Javkin, Hector R. 1976. "The Perceptual Basis of Vowel Duration Differences Associated with the Voiced/Voiceless Distinction." *Report of the Phonology Laboratory, UC Berkeley* 1: 78–92.
- Jong, Kenneth de. 1991. "An Articulatory Study of Consonant-Induced Vowel Duration Changes in English." *Phonetica* 48 (1): 1–17.
- . 2004. "Stress, Lexical Focus, and Segmental Focus in English: Patterns of Variation in Vowel Duration." *Journal of Phonetics* 32 (4): 493–516.
- Kirby, James, and Morgan Sonderegger. 2018. "Mixed-Effects Design Analysis for Experimental Phonetics." *Journal of Phonetics* 70: 70–85.
- Klatt, Dennis H. 1973. "Interaction Between Two Factors That Influence Vowel Duration." *The Journal of the Acoustical Society of America* 54 (4): 1102–4.
- Kluender, Keith R., Randy L. Diehl, and Beverly A. Wright. 1988. "Vowel-Length Differences Before Voiced and Voiceless Consonants: An Auditory Explanation." *Journal of Phonetics* 16: 153–69.
- Ko, Eon-Suk. 2018. "Asymmetric Effects of Speaking Rate on the Vowel/Consonant Ratio Conditioned by Coda Voicing in English." *Phonetics and Speech Sciences* 10 (2): 45–50.
- Laeuffer, Christiane. 1992. "Patterns of Voicing-Conditioned Vowel Duration in French and English." *Journal of Phonetics* 20 (4): 411–40.
- Lau, Joseph, John P. A. Ioannidis, Norma Terrin, Christopher H. Schmid, and Ingram Olkin. 2006. "The Case of the Misleading Funnel Plot." *British Medical Journal* 333 (7568): 597–600.
- Light, Richard J., and David B. Pillemer. 1984. *Summing up; the Science of Reviewing Research*. Cambridge, MA (USA) Harvard Univ. Press.

- Lindblom, Björn. 1967. "Vowel Duration and a Model of Lip Mandible Coordination." *Speech Transmission Laboratory Quarterly Progress Status Report* 4: 1–29.
- Lisker, Leigh. 1957. "Closure Duration and the Intervocalic Voiced-Voiceless Distinction in English." *Language* 33 (1): 42–49.
- . 1974. "On 'Explaining' Vowel Duration Variation." In *Proceedings of the Linguistic Society of America*, 225–32.
- Luce, Paul A., and Jan Charles-Luce. 1985. "Contextual Effects on Vowel Duration, Closure Duration, and the Consonant/Vowel Ratio in Speech Production." *The Journal of the Acoustical Society of America* 78 (6): 1949–57.
- Mack, Molly. 1982. "Voicing-dependent Vowel Duration in English and French: Monolingual and Bilingual Production." *The Journal of the Acoustical Society of America* 71 (1): 173–78.
- Maddieson, Ian, and Jack Gandour. 1976. "Vowel Length Before Aspirated Consonants." In *UCLA Working papers in Phonetics*, 31:46–52.
- Meyer, Ernst Alfred. 1903. *Englische Lautdauer: Eine Experimentalphonetische Untersuchung*. Akademiska Bokhandeln, Uppsala, Sweden.
- . 1904. "Zur Vokaldauer Im Deutschen." In *Nordiska Studier Tillegnade A. Noreen*, 347–56. Uppsala: K.W. Appelbergs Boktryckeri.
- Meyer, Ernst Alfred, and Z. Gombocz. 1909. "Zur Phonetik Der Ungarischen Sprache." *Le Monde Oriental* 3: 122–97.
- Mitleb, Fares. 1982. "Voicing Effect on Vowel Duration Is Not an Absolute Universal." *The Journal of the Acoustical Society of America* 71 (S1): S23–S23.
- Nicenboim, Bruno, Timo B. Roettger, and Shravan Vasisht. 2018. "Using Meta-Analysis for Evidence Synthesis: The Case of Incomplete Neutralization in German." *Journal of Phonetics* 70: 39–55.
- Peterson, Gordon E., and Ilse Lehiste. 1960. "Duration of Syllable Nuclei in English." *The Journal of the Acoustical Society of America* 32 (6): 693–703.
- Port, Robert F., and Jonathan Dalby. 1982. "Consonant/Vowel Ratio as a Cue for Voicing in English." *Perception & Psychophysics* 32 (2): 141–52.
- Raphael, Lawrence J. 1975. "The Physiological Control of Durational Differences Between Vowels Preceding Voiced and Voiceless Consonants in English." *Journal of Phonetics* 3 (1): 25–33.
- Ratnikova, E. I. 2017. "Towards a Log-Normal Model of Phonation Units Lengths Distribution in the Oral Utterances." *International Research Journal* 3 (57): 46–49.
- Roettger, Timo B., Bodo Winter, and Harald Baayen. 2018. "Emergent Data Analysis in Phonetic Sciences: Towards Pluralism and Reproducibility." *Journal of Phonetics* 73: 1–7.
- Rosen, Kristin M. 2005. "Analysis of Speech Segment Duration with the Lognormal Distribution: A Basis for Unification and Comparison." *Journal of Phonetics* 33 (4): 411–26.
- Sanker, Chelsea. 2019. "Influence of Coda Stop Features on Perceived Vowel Duration." *Journal of Phonetics* 75: 43–56.
- Sharf, Donald J. 1962. "Duration of Post-Stress Intervocalic Stops and Preceding Vowels." *Language and Speech* 5 (1): 26–30.
- . 1964. "Vowel Duration in Whispered and in Normal Speech." *Language and Speech* 7 (2): 89–97.
- Slis, Iman Hans, and Antonie Cohen. 1969a. "On the Complex Regulating the Voiced-Voiceless Distinction I." *Language and Speech* 12 (2): 80–102.

- . 1969b. “On the Complex Regulating the Voiced-Voiceless Distinction II.” *Language and Speech* 12 (3): 137–55.
- Sterne, Jonathan A. C., Betsy Jane Becker, and Matthias Egger. 2005. “The Funnel Plot.” In *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, edited by Hannah R. Rothstein, Alexander J. Sutton, and Michael Borenstein, 75–98. Wiley Online Library.
- Summers, W. Van. 1987. “Effects of Stress and Final-Consonant Voicing on Vowel Production: Articulatory and Acoustic Analyses.” *The Journal of the Acoustical Society of America* 82 (3): 847–63.
- Umeda, Noriko. 1975. “Vowel Duration in American English.” *The Journal of the Acoustical Society of America* 58 (2): 434–45.
- Vasishth, Shravan, and Andrew Gelman. 2019. “How to Embrace Variation and Accept Uncertainty in Linguistic and Psycholinguistic Data.” PsyArXiv.
- Warren, Willis, and Adam Jacks. 2005. “Lip and Jaw Closing Gesture Durations in Syllable Final Voiced and Voiceless Stops.” *The Journal of the Acoustical Society of America* 117 (4): 2618–8.
- Wells, John C. 1990. “Syllabification and Allophony.” In *Studies in the Pronunciation of English: A Commemorative Volume in Honour of A. C. Gimson*, edited by Susan Ramsaran, 76–86. New York: Routledge.
- Zimmerman, Samuel A., and Stanley M. Sapon. 1958. “Note on Vowel Duration Seen Cross-Linguistically.” *The Journal of the Acoustical Society of America* 30 (2): 152–53.