Bayesian meta-analysis of the voicing effect in English

Stefano Coretta

A Bayesian meta-analysis of the English voicing effect has been run on the basis of 11 estimated posterior distributions extracted from 9 different publications, following the procedures discussed in Nicenboim et al. (2018). The studies were selected by scraping the first 100 results on Google Scholar with the keywords "vowel duration voicing English". Other studies which were known to the author but not among the Google Scholar results were also included.

Since two publications, Sharf (1962) and Klatt (1973), tested both monosyllabic and disyllabic words, two separate posterior distributions were estimated for each word type. This leads to a total of 11 posterior distribution of the effect of consonant voicing on vowel duration in English (7 estimated posteriors from 7 publications plus 2 each from 2 publications).

The posterior distributions of each study have been obtained by fitting a Bayesian linear model to the summary data (the means of vowel duration before voiceless and voiced stops) provided by the respective publications. These models had the mean vowel durations as outcome and consonant voicing (voiceless vs. voiced) as the only predictor. Three studies, Luce & Charles-Luce (1985), Davis & Van Summers (1989), and Ko (2018), reported measures of dispersion along with the means. Measurement error models were used to obtain the posterior distributions from these studies. The measurement error term in such models allows fitting to include information of the dispersion of the mean vowel durations, and hence of the uncertainty that comes with them. All the models for estimating the posterior of the individual studies were fitted with the following priors: a normal distribution with mean = 0 ms and SD = 300 for the intercept, and a normal distribution with mean = 0 ms and SD = 100 for the effect of consonant voicing. The simple models (without an error term) also included a prior for the residual variance as a half Cauchy distribution with location = 0 ms and scale = 25.

A data set with the mean estimates and estimated standard errors from these 11 posterior distributions has then been used to fit a further Bayesian measurement error model. In this model, the mean estimates with the estimated standard errors were included as the outcome, while a by-study random intercept was the only predictor.

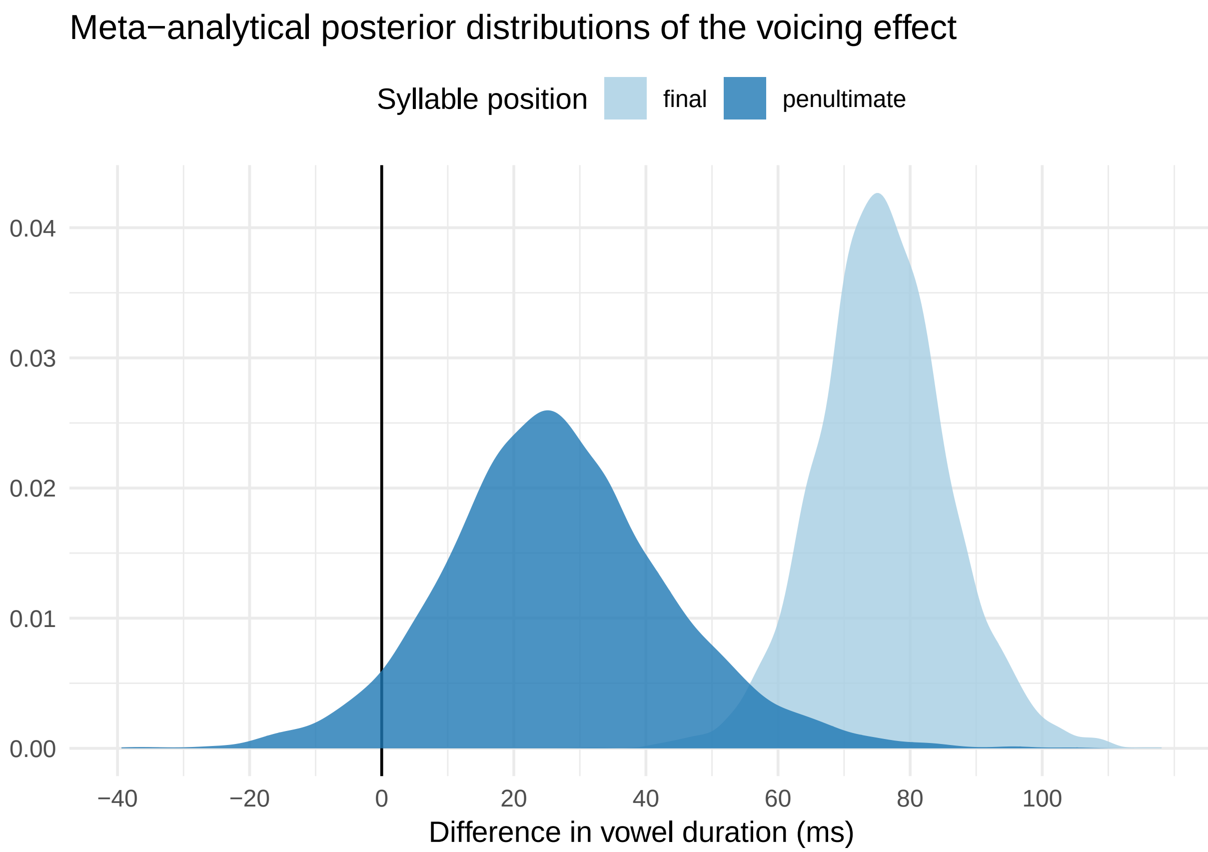
The following is the summary of the meta-analytical model.

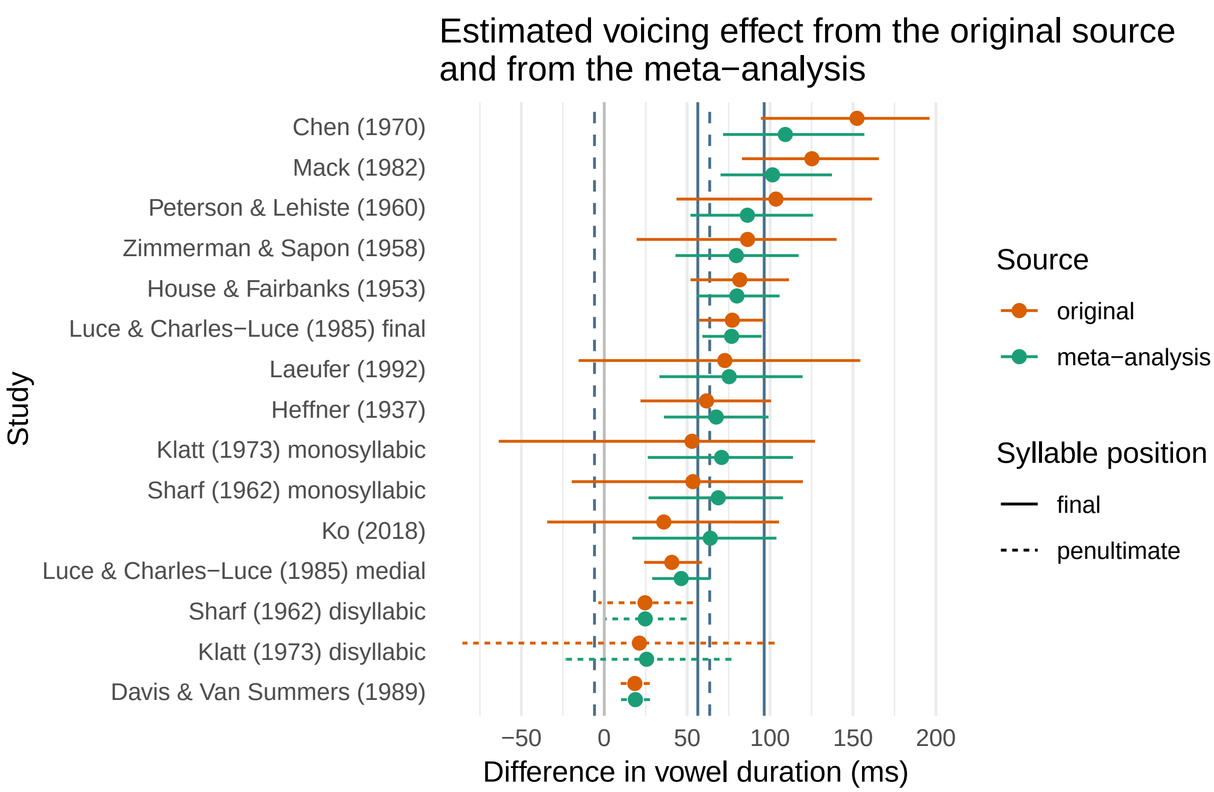
The 95% credible interval (CI) of the model intercept posterior (which corresponds to the estimated voicing effect in word-final syllables) is between 56.39 and 96.43 ms. The mean of the posterior distribution is 75.83 ms (estimated error = 10.01). Given the posterior and based on the surveyed studies, it is safe to infer that the true effect of voicing in word-final syllables in English is very likely less than 100 ms.

The posterior mean of the coefficient when the target syllable is in penultimate position is -49.14 ms (est.e = 19.10). Note that the estimated error is double compared to that of the intercept, which increases uncertainty. On average, we can argue that the mean voicing effect in penultimate syllables is about 50 ms less than the mean effect in monosyllabic words. The mean of the voicing effect in disyllabic words can thus be estimated to be around 25 ms. **??** shows the full posterior distributions of the voicing effect in the word-final and penultimate contexts. Note how the calculated posterior distribution in penultimate position has greater dispersion that the other.

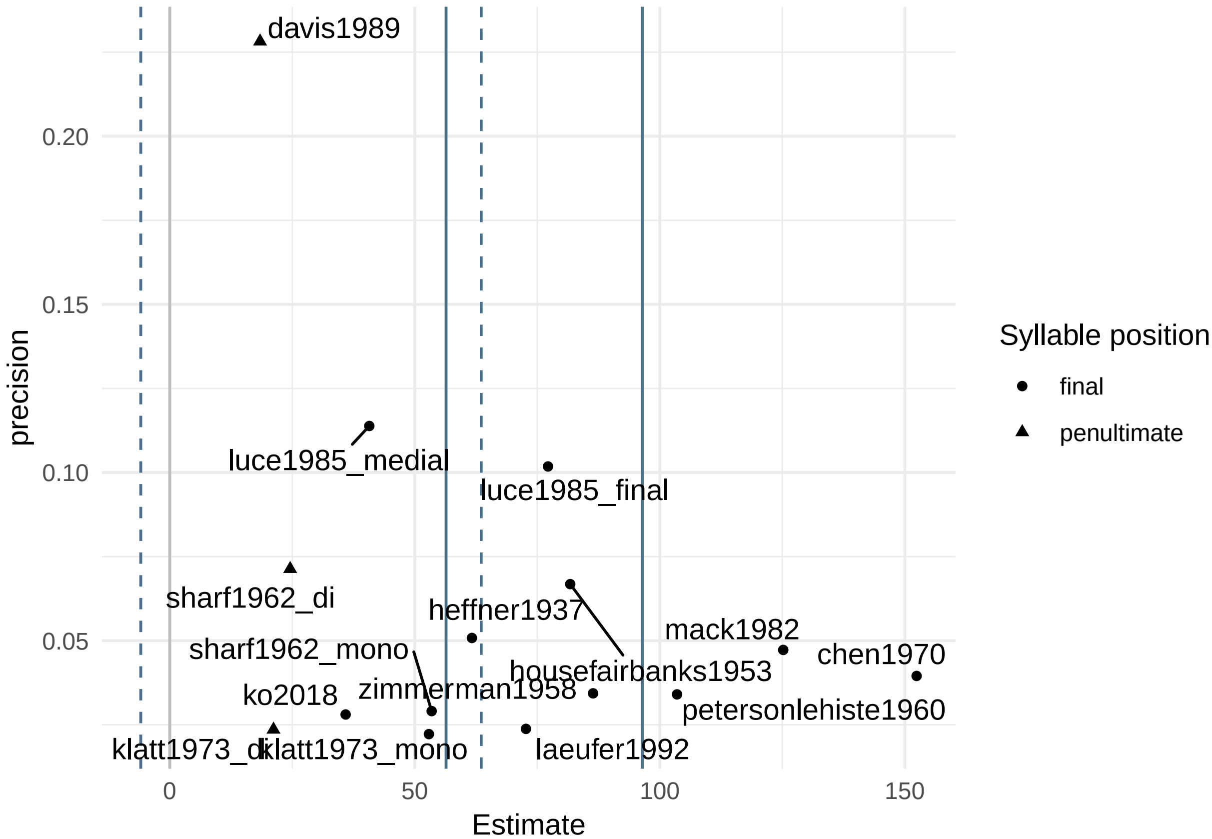
?? shows the mean estimate (the points) of the voicing effect with 95% CIs (the horizontal segments) for each of the 11 studies. For each study, the plot gives both the original estimate (as obtained from the raw data summary of the study) and the estimate shrunk by the random effects in the meta-analytical model. The vertical lines indicate the meta-analytical 95% CI of the voicing effect in final (solid) and penultimate syllable position (dashed). Original estimates further away from the meta-analytical mean effect and those with greater uncertainty (wider errors) show greater shrinkage to the mean.

A funnel plot can be used to visually check whether the sample suffers from publication bias. In **??**, the x-axis corresponds to the original estimated difference in vowel duration, while the y-axis is a measure of precision (calculated as 1 divided by the estimated error of the difference). When there is no bias, the points with lower precision should be more spread out and symmetrical placed around the meta-analytical mean, while points with higher precision should cluster around the mean. This ideal situation is clearly not the case for the final syllable context. There





The vertical lines indicate the 95% CI of the meta−analytical posterior of the voicing effect in final (solid) and penultimate (dashed) position.



seems to be a bias towards bigger effects (which also happen to have lower precision).

References

Davis, Stuart & W. Van Summers. 1989. Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics* 17. 339–353.

Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America* 54(4). 1102–1104. doi:10.1121/1.1914322.

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. doi:10.13064/KSSS.2018.10.2.045.

Luce, Paul A & 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). 1949–1957.

Nicenboim, Bruno, Timo B. Roettger & Shravan Vasishth. 2018. Using meta-analysis for evidence synthesis: The case of incomplete neutralization in german. *Journal of Phonetics* 70. 39–55. doi:10.1016/j.wocn.2018.06.001.

Sharf, Donald J. 1962. Duration of post-stress intervocalic stops and preceding vowels. *Language and speech* 5(1). 26–30.