

PraatSauce/VoiceSauce comparison: Madurese

James Kirby

31/01/2018

Contents

Data	1
Plots	2
F0	2
Formants	3
Bandwidths	5
Uncorrected amplitudes	8
Corrected amplitudes	10
Corrected differences	24
Cepstral peak prominence	28
Harmonic to noise ratios	29
Distinguishing vowel qualities	30
High/back vowels	30
High/front vowels	33
Low/front vowels	36

Data

This small dataset compares spectral measures generated by both PraatSauce v0.2.2 and VoiceSauce v1.31 at 1 msec intervals for 9 Madurese lexical items spoken by a single male speaker. (There are 12 items included, but here we set aside the centralised vowels for the time being.) The original audio files, included here, are from Misnadin and Kirby (forthcoming). For both scripts, 5 formants were estimated with a maximum formant frequency of 5000 Hz; minimum and maximum F0 values were set to 50 Hz and 300 Hz for all F0 estimators. For VoiceSauce, the STRAIGHT F0 estimate and Snack formant/bandwidth estimates were used for harmonic amplitude corrections.

The `method` column indicates whether the formant bandwidths were `estimated` using Praat (PraatSauce) or Snack (VoiceSauce), or whether the Hawks and Miller `formula` was used.

Madurese has a consonant-vowel co-occurrence restriction whereby vowels are organized into high/non-high pairs. The high member of each pair follows voiced and voiceless aspirated stops, while the non-high member follows the voiceless unaspirated member. The items analysed here represent exemplars of each vowel quality with a bilabial onset, in order to explore the effects of the spectral harmonic correction implementations.

```
head(df)
```

```
##      Filename Item Gloss seg_Start seg_End   t_ms          t method
## 1 baca-read baca  read   99.119 200.695  99.119 0.00000000 formula
## 2 baca-read baca  read   99.119 200.695 100.119 0.00990099 formula
## 3 baca-read baca  read   99.119 200.695 101.119 0.01980198 formula
## 4 baca-read baca  read   99.119 200.695 102.119 0.02970297 formula
## 5 baca-read baca  read   99.119 200.695 103.119 0.03960396 formula
## 6 baca-read baca  read   99.119 200.695 104.119 0.04950495 formula
##          script measure  value  corrected
## 1 PraatSauce      pF0 126.646 uncorrected
## 2 PraatSauce      pF0 126.571 uncorrected
## 3 PraatSauce      pF0 126.497 uncorrected
```

```
## 4 PraatSauce      pF0 126.422 uncorrected
## 5 PraatSauce      pF0 126.347 uncorrected
## 6 PraatSauce      pF0 126.273 uncorrected
```

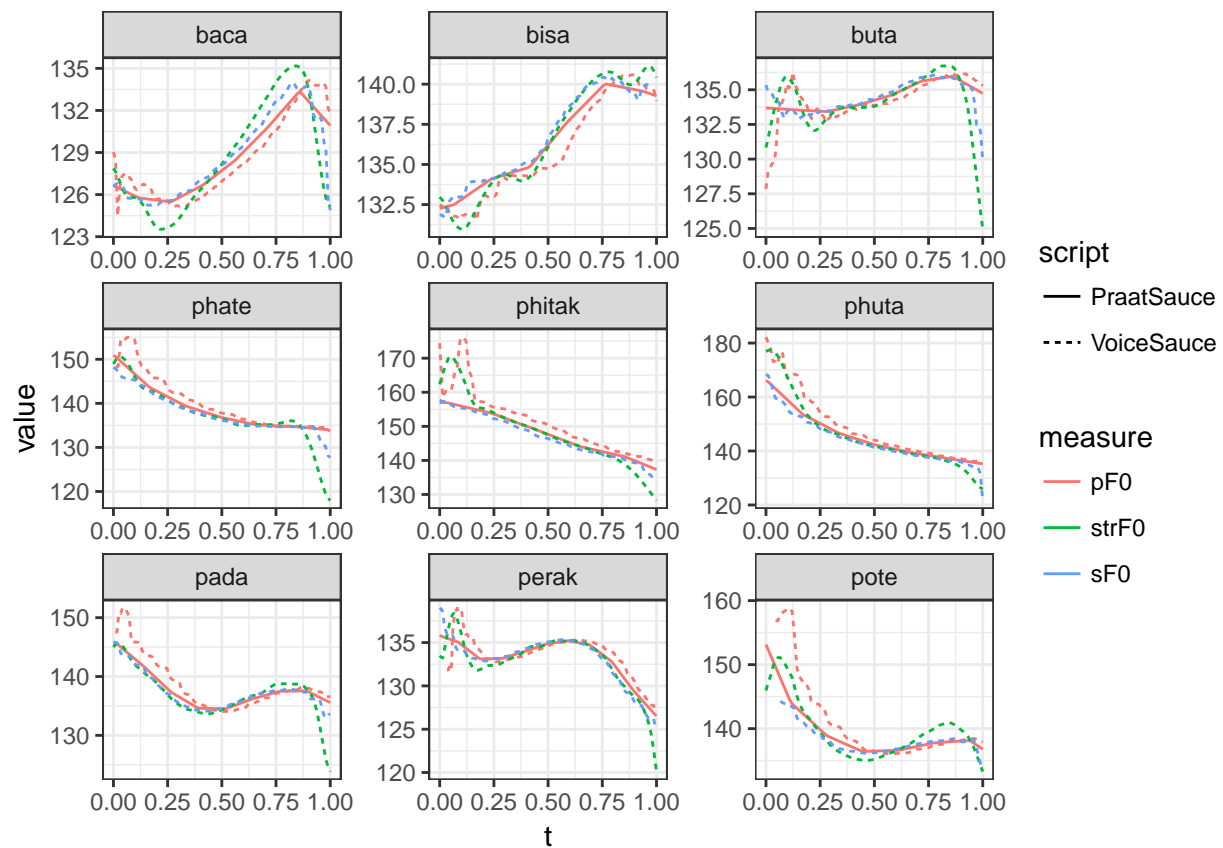
In the plots which follow, the PraatSauce measures are unsmoothed. If you want to compare to smoothed estimates, uncomment the two lines:

```
ps.fbw <- cbind(ps.fbw[1:6], apply(ps.fbw[7:43], 2, filter, filter=f21, sides=2))
ps.ebw <- cbind(ps.ebw[1:6], apply(ps.ebw[7:43], 2, filter, filter=f21, sides=2))
```

If you want to smooth the Matlab way, use the lag kernel by selecting `filter=f20` and set `sides=1`.

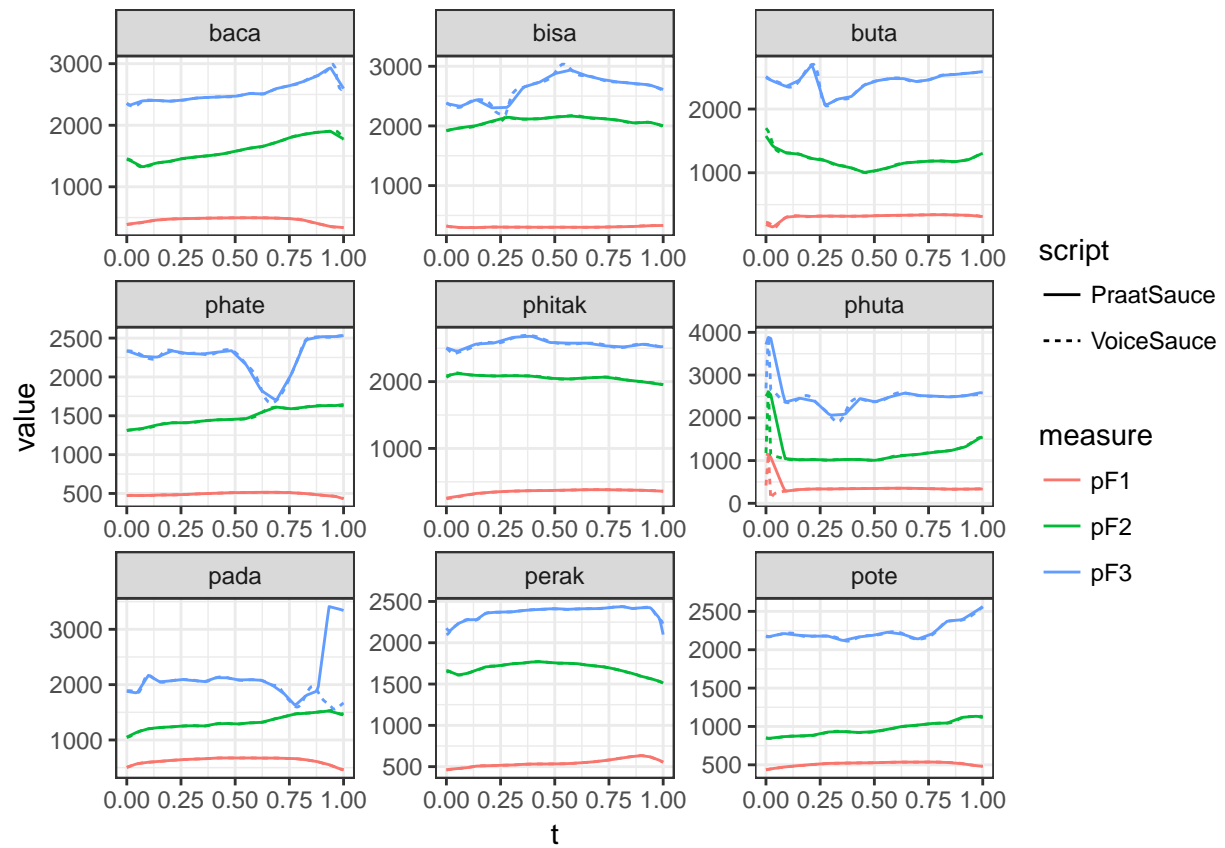
Plots

F0

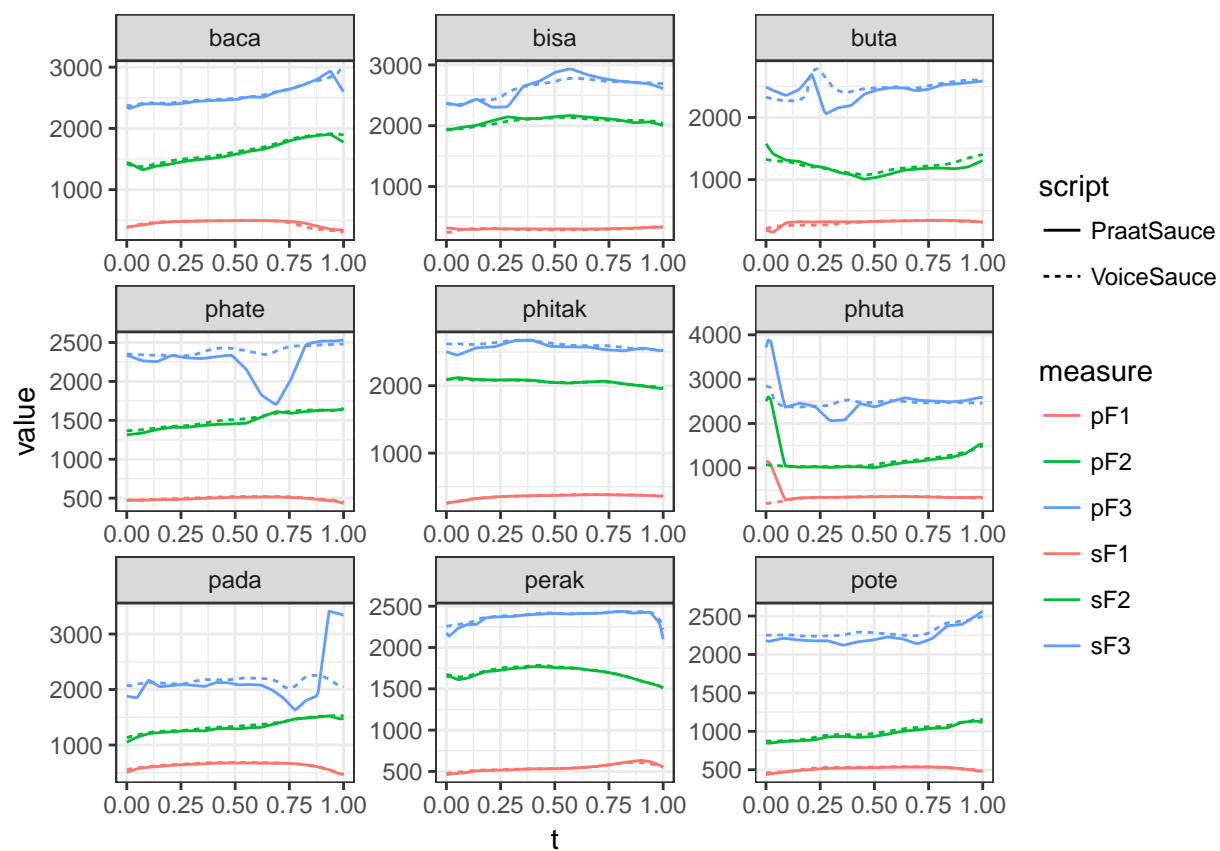


It is not clear why VoiceSauce's pF0 (Praat) estimate differs so dramatically from the PraatSauce estimate, considering they should be using the same estimator with the same settings.

Formants

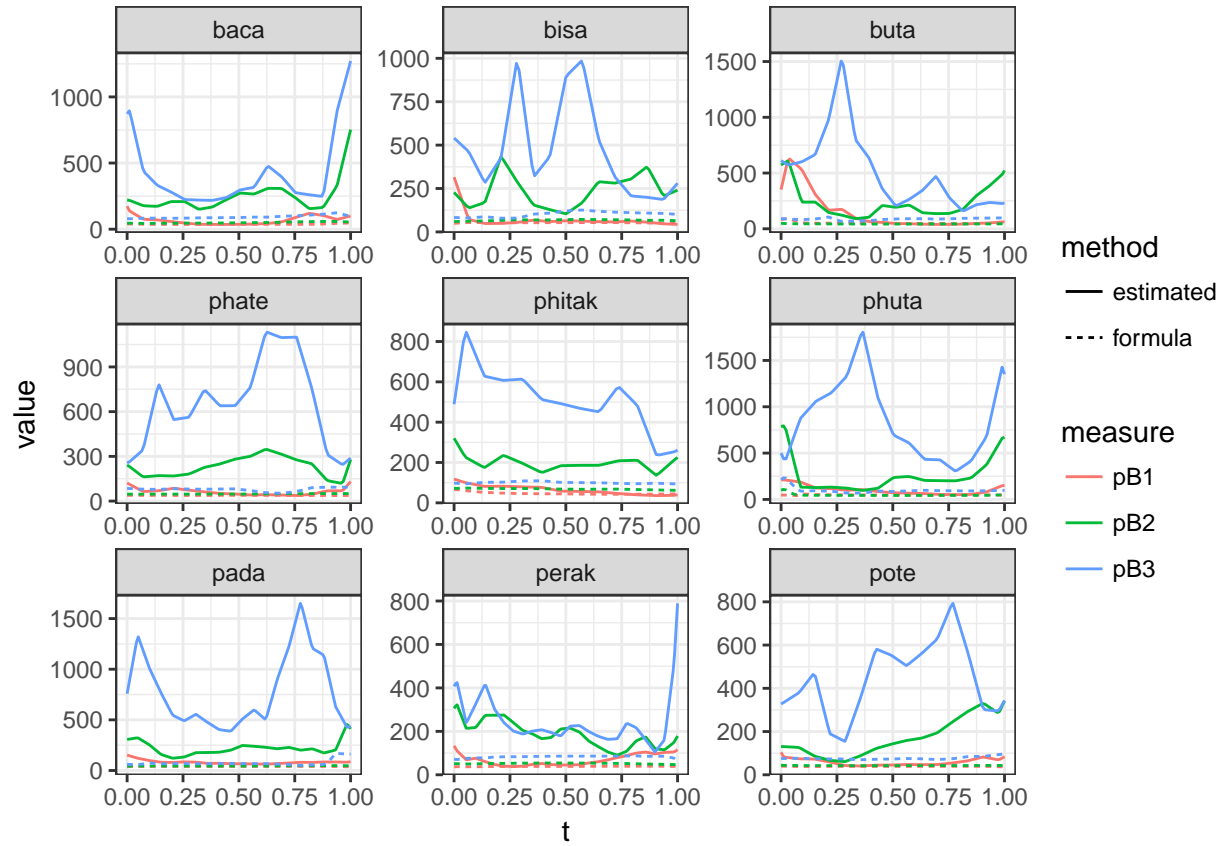


Compare with the VoiceSauce Snack estimates:



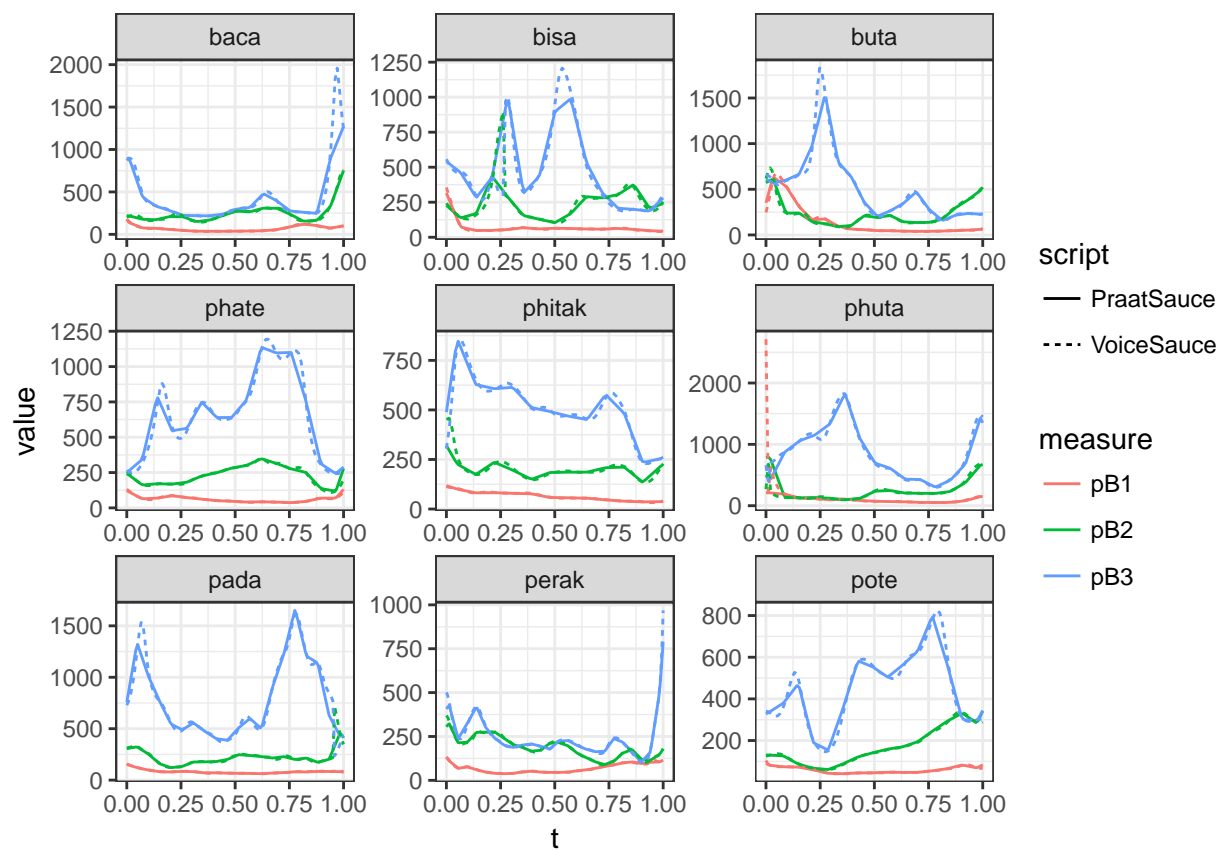
Bandwidths

PraatSauce estimated vs. formula bandwidths



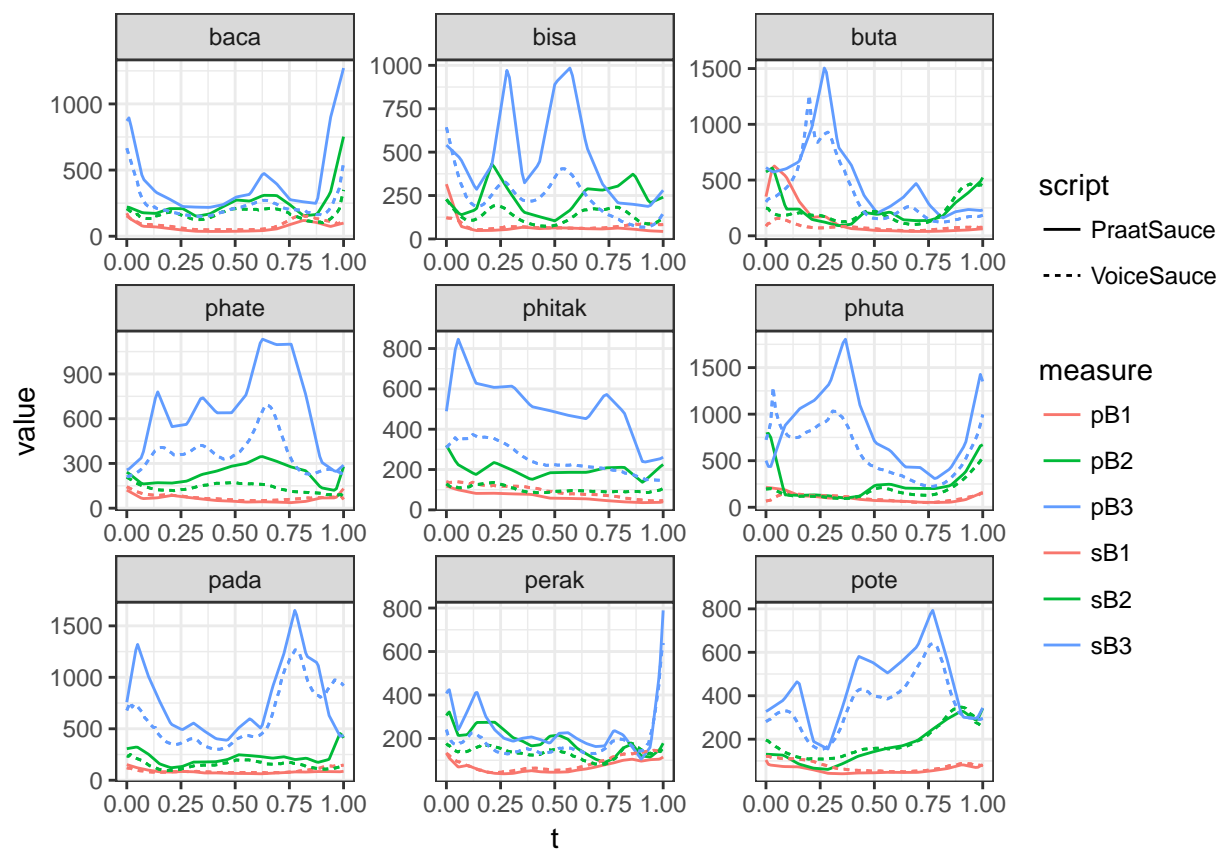
PraatSauce estimated bandwidths are huge...

PraatSauce vs. VoiceSauce estimated bandwidths



Here, both Praat-based estimators seem to be sync.

VoiceSauce Snack vs. PraatSauce estimated bandwidths

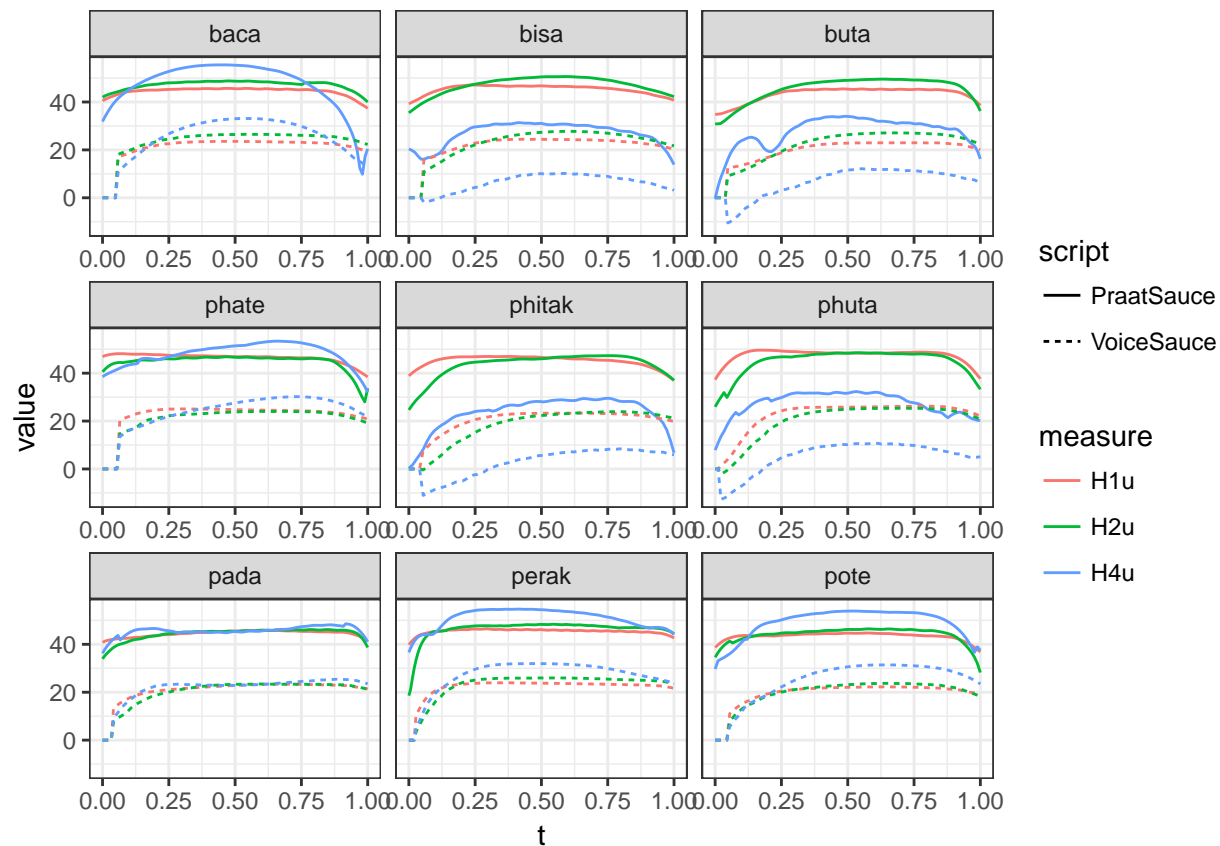


PraatSauce estimates not completely off from Snack (if they really are Snack estimates).

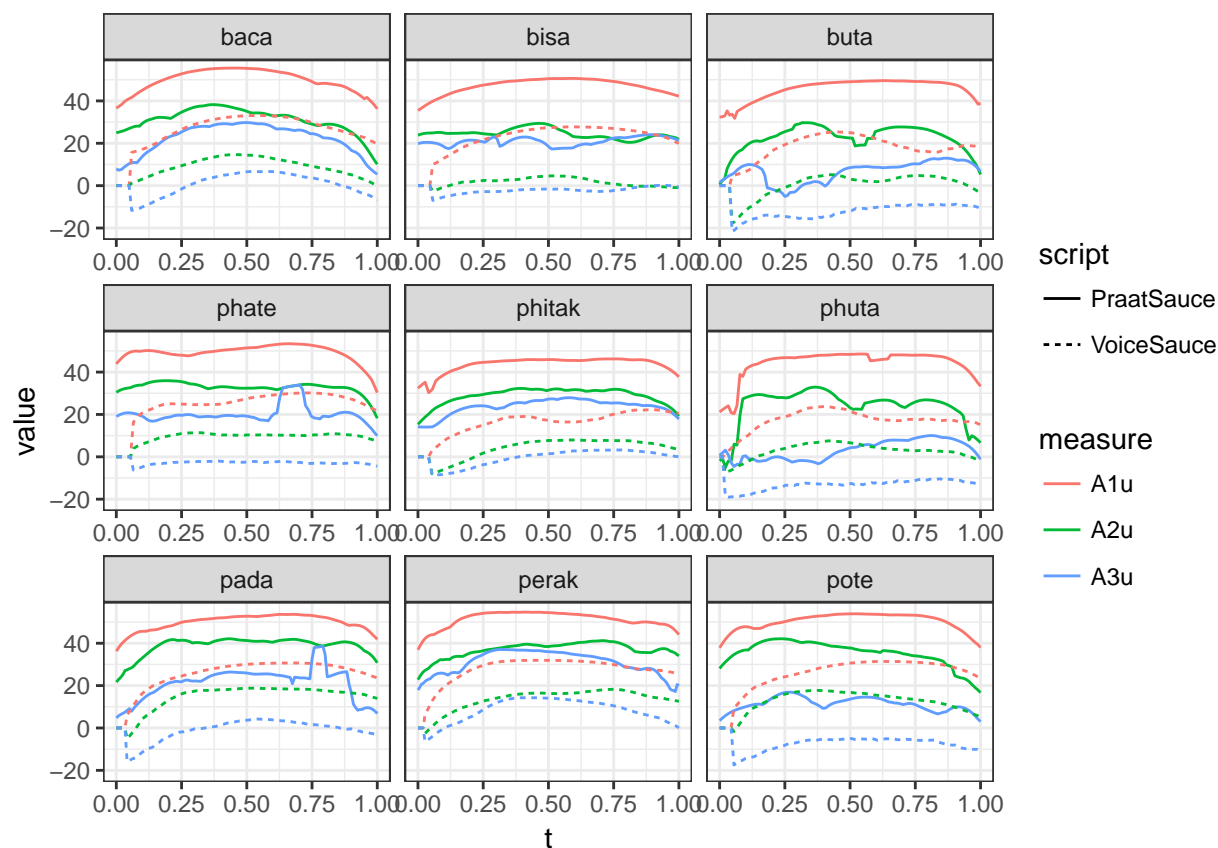
Uncorrected amplitudes

PraatSauce vs. VoiceSauce H1, H2, H4

Note that the choice of bandwidth estimator is irrelevant here.



PraatSauce vs. VoiceSauce A1-A3



For reasons I have not been able to work out, PraatSauce estimates are consistently 20-25 dB higher than VoiceSauce estimates. This doesn't matter for the spectral differences, but it would be good to work out what the source of the difference is – probably some kind of amplitude normalization/attenuation being done by VoiceSauce somewhere.

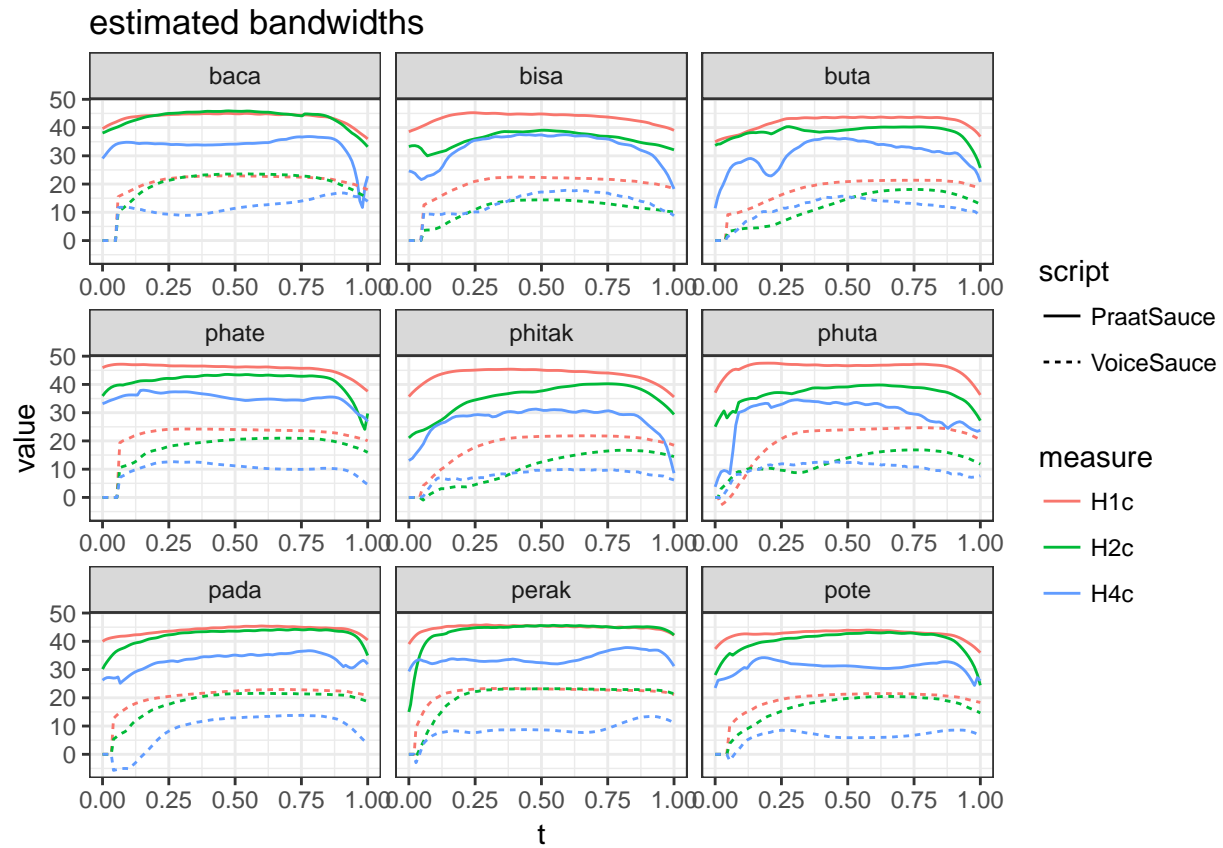
Note that the VoiceSauce estimates are sometimes negative, which seems...strange.

Corrected amplitudes

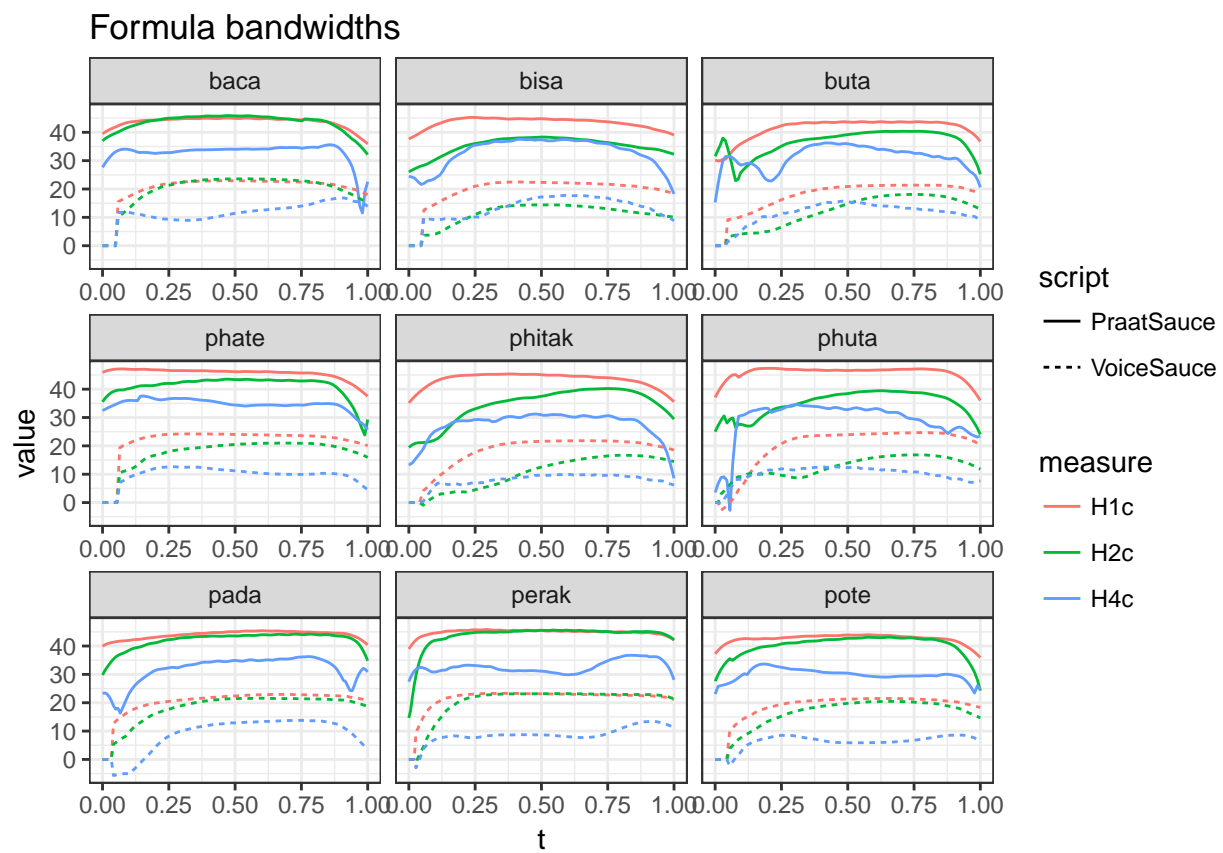
Here, choice of formant bandwidth estimator potentially matters.

In these plots, PraatSauce is using Praat and VoiceSauce is using Snack estimates.

PraatSauce vs. VoiceSauce H1*, H2*, H4*

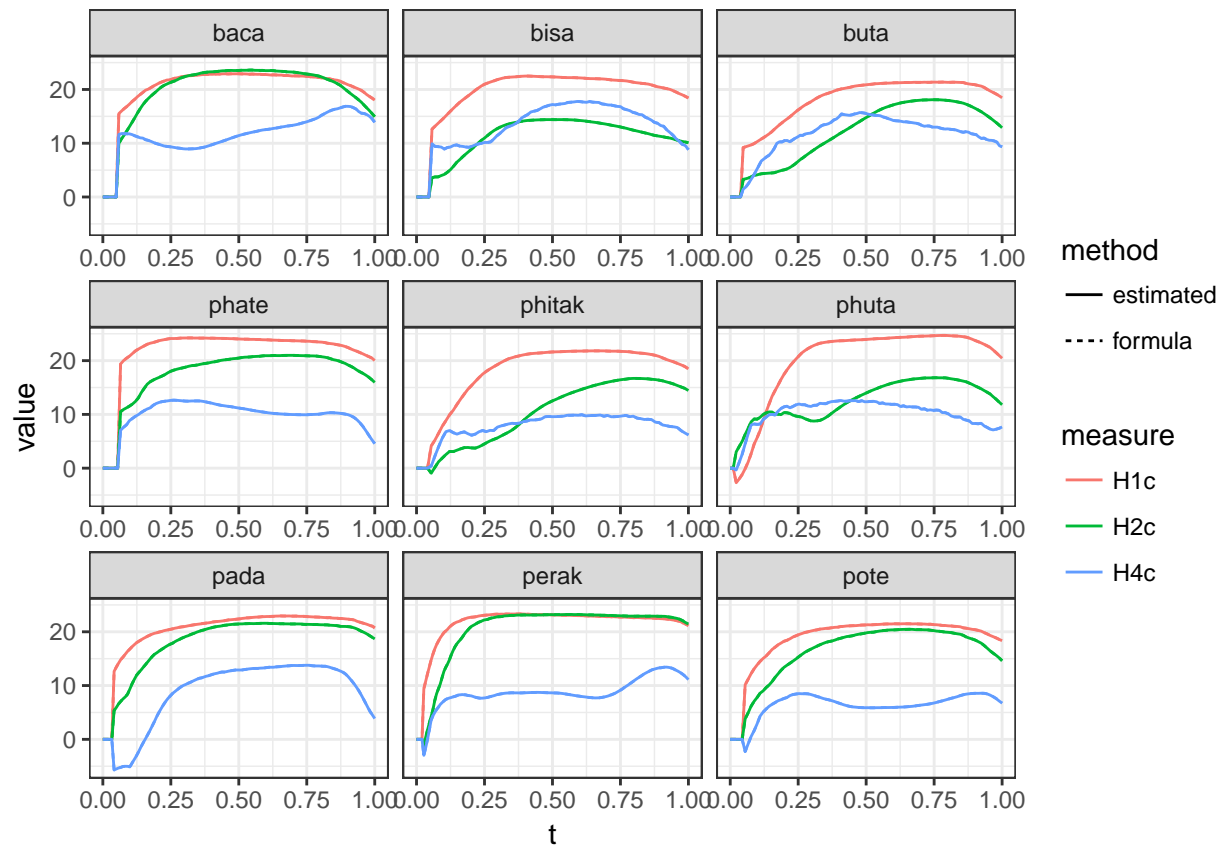


PraatSauce vs. VoiceSauce H1*, H2*, H4*



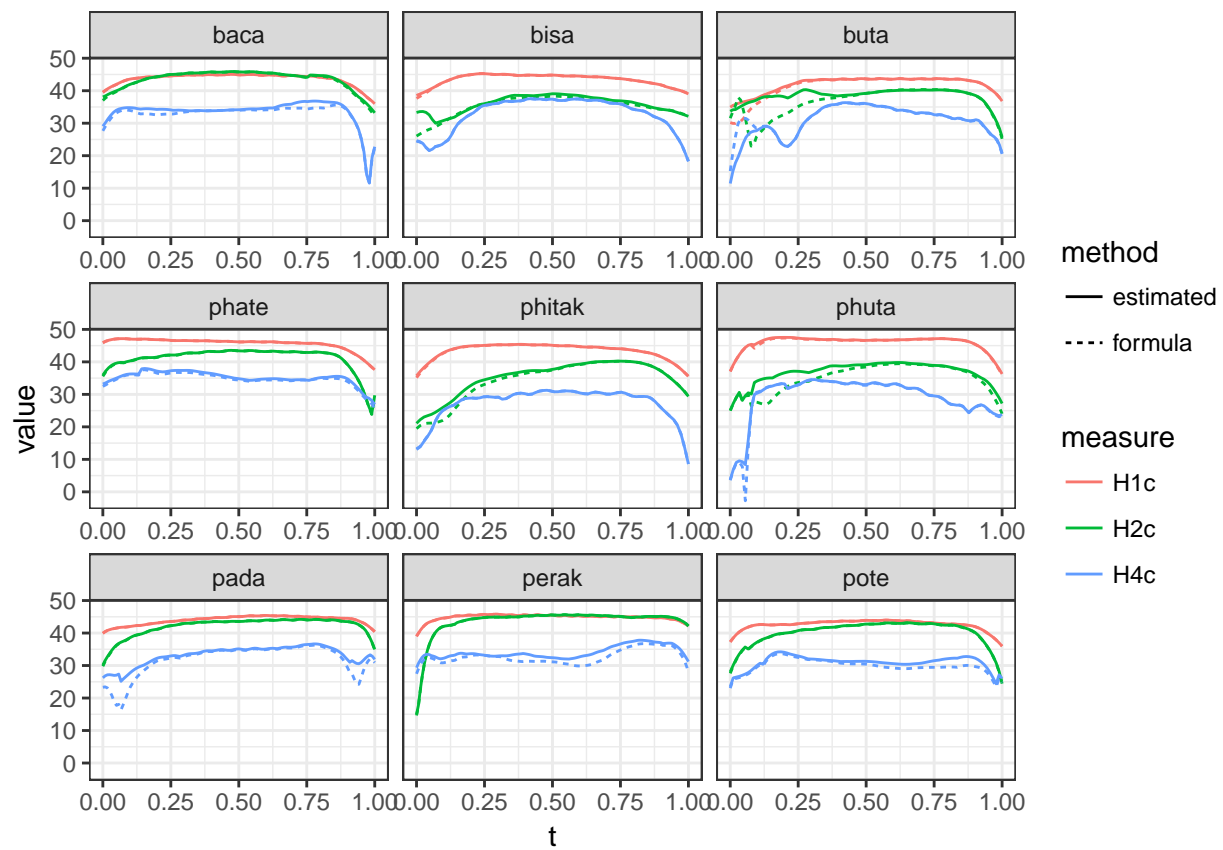
For VoiceSauce, no real difference can be observed:

VoiceSauce estimated vs. formula bandwidths, H1*, H2*, H4*

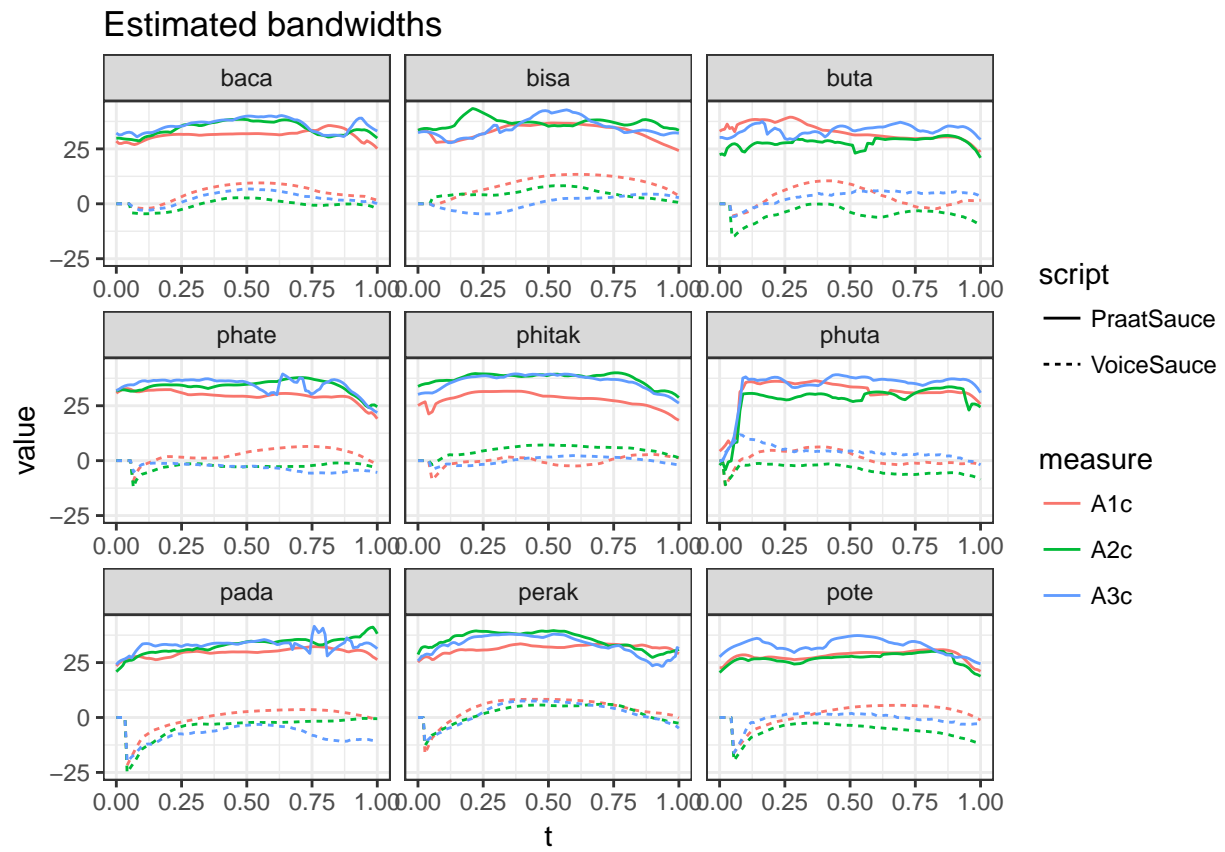


For PraatSauce, using the formula bandwidths makes only minor differences:

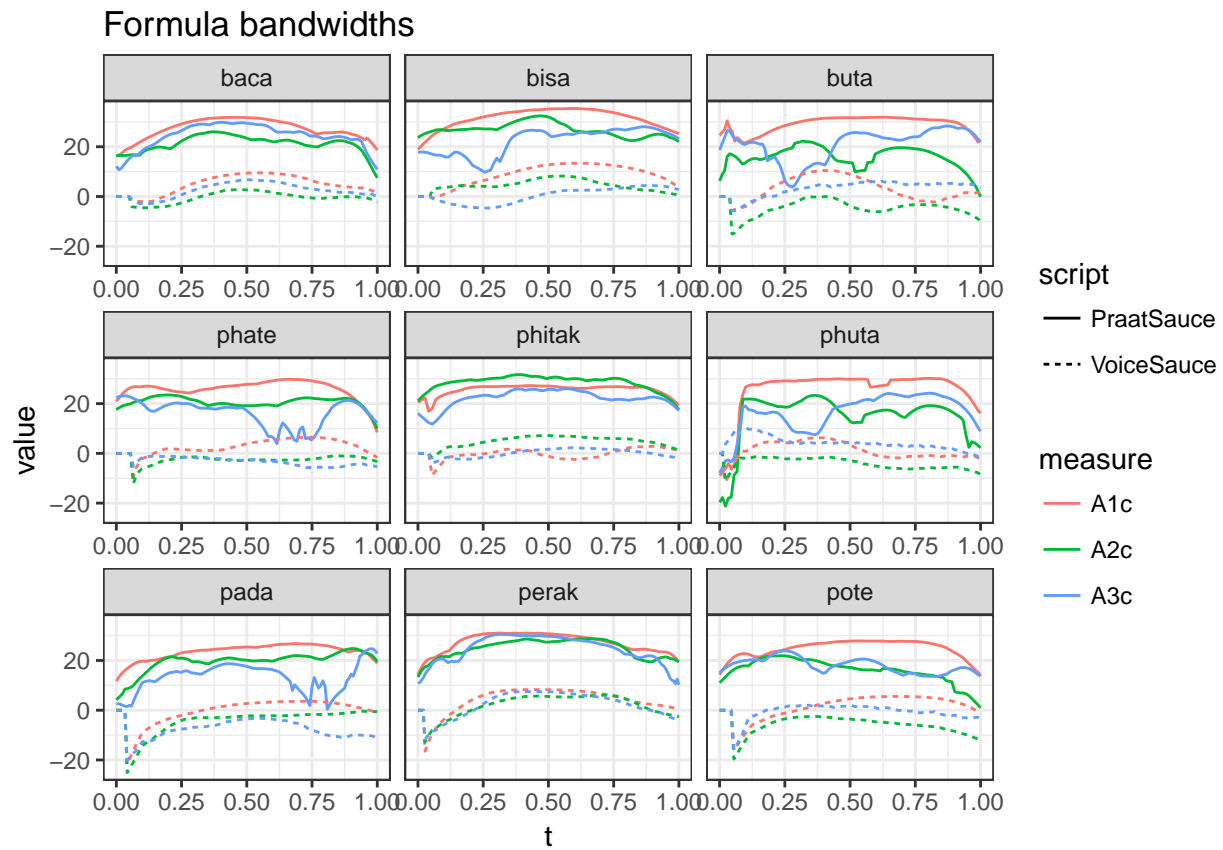
PraatSauce estimated vs. formula bandwidths, H1*, H2*, H4*



PraatSauce vs. VoiceSauce A1*, A2*, A3*

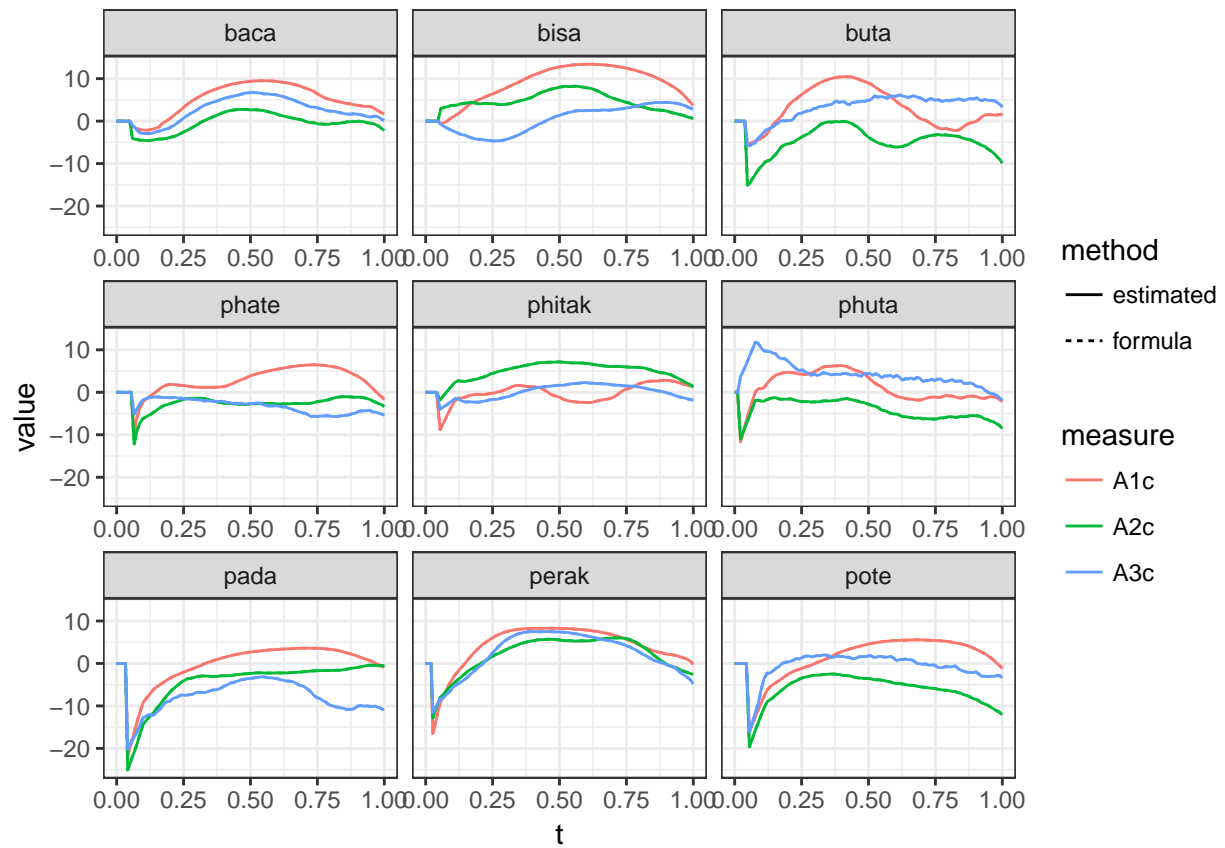


PraatSauce vs. VoiceSauce A1*, A2*, A3*



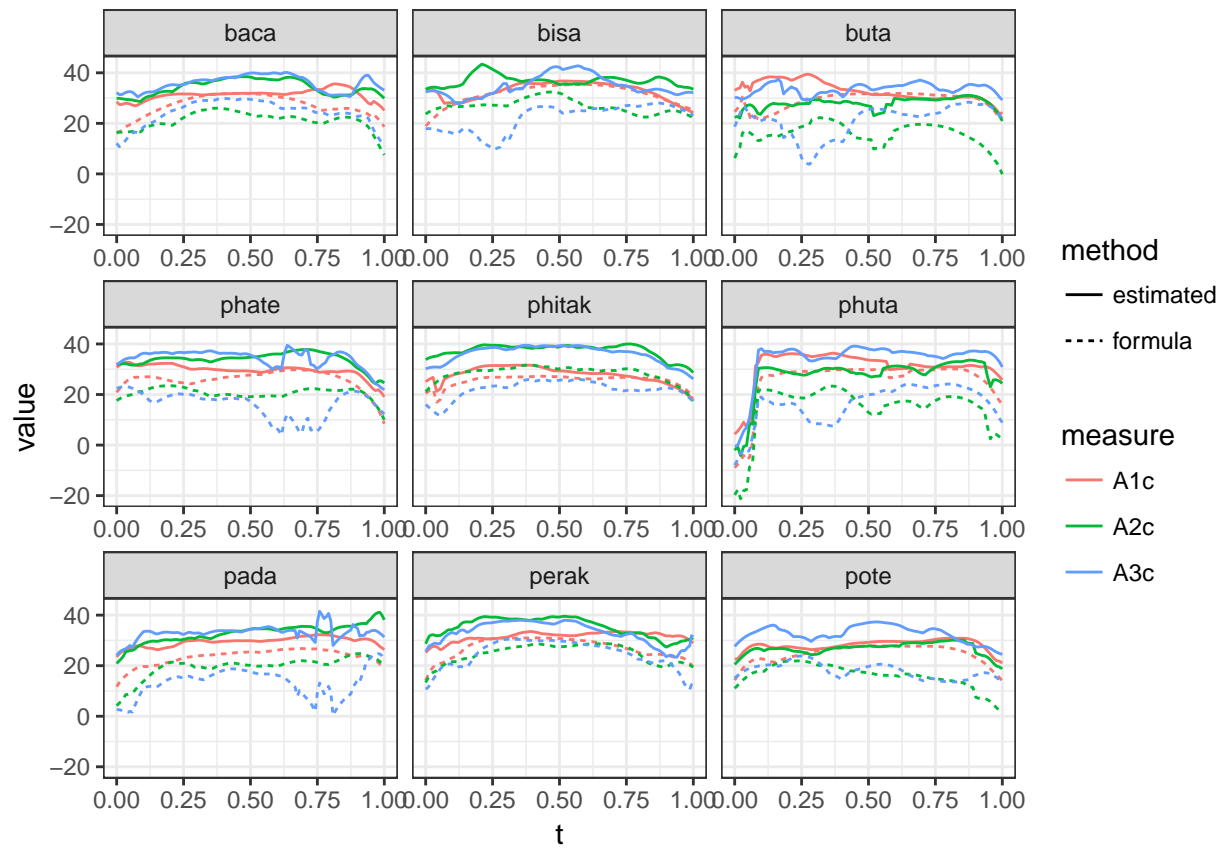
Again for VS the choice of bandwidth estimate doesn't seem to matter:

VoiceSauce estimated vs. formula bandwidths, A1*, A2*, A3*

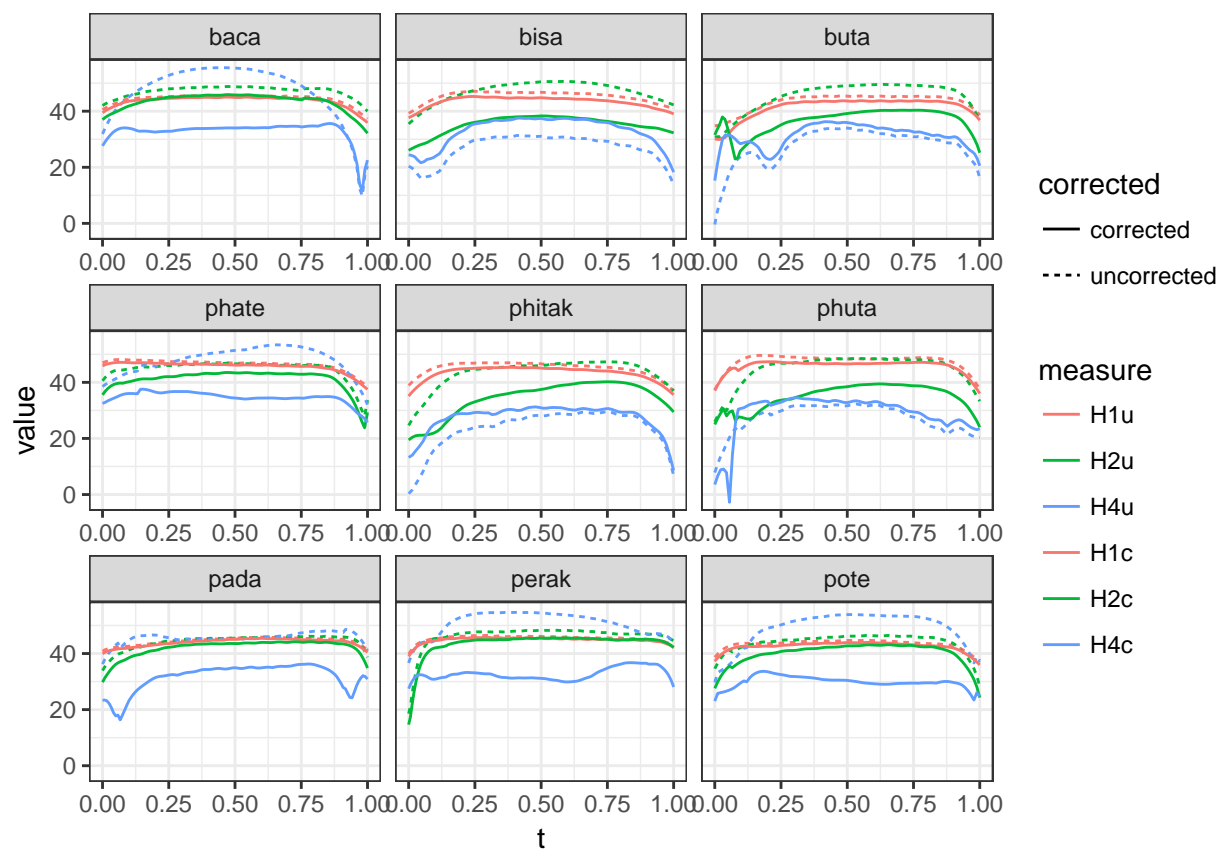


For PraatSauce things are less clear:

PraatSauce estimated vs. formula bandwidths, A1*, A2*, A3*

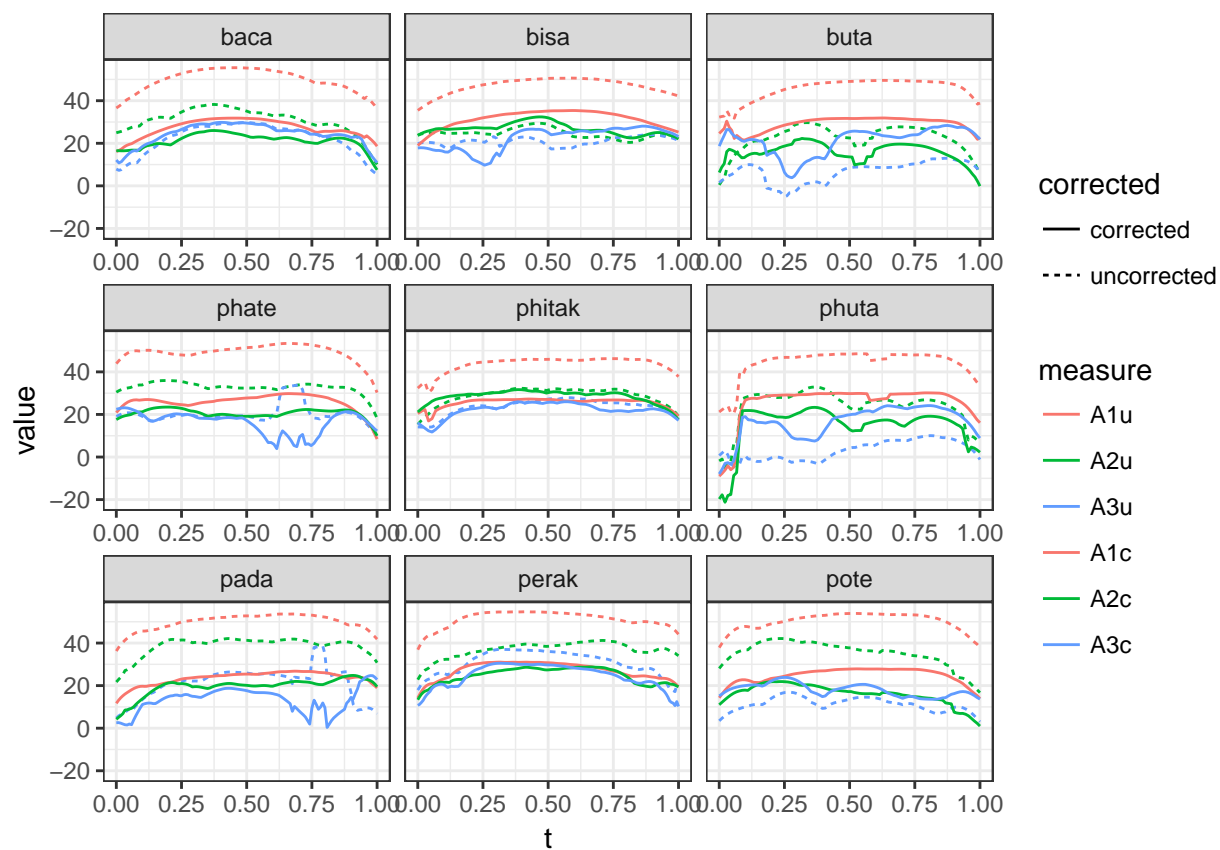


PraatSauce corrected vs. uncorrected, H1/H2/H4, formula bandwidths



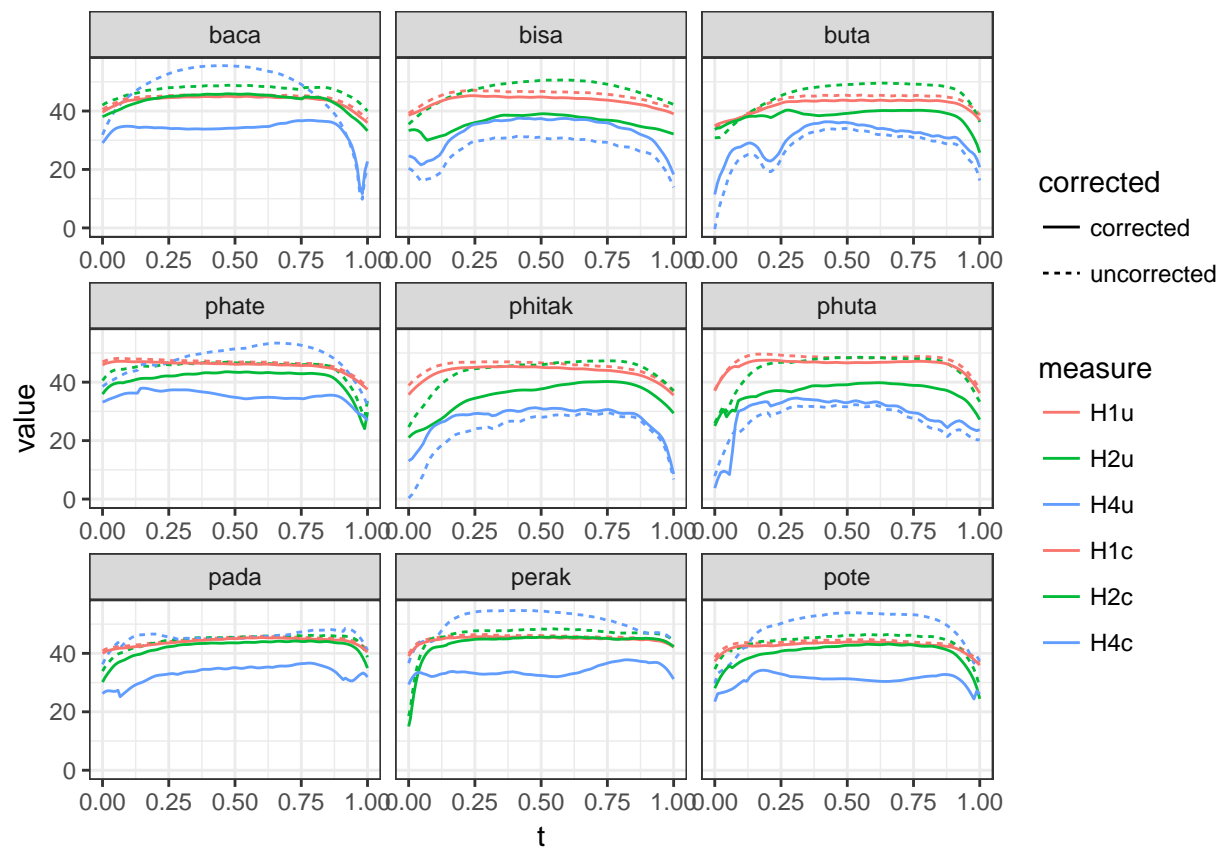
Largest difference for H4.

PraatSauce corrected vs. uncorrected, A1/A2/A3, formula bandwidths

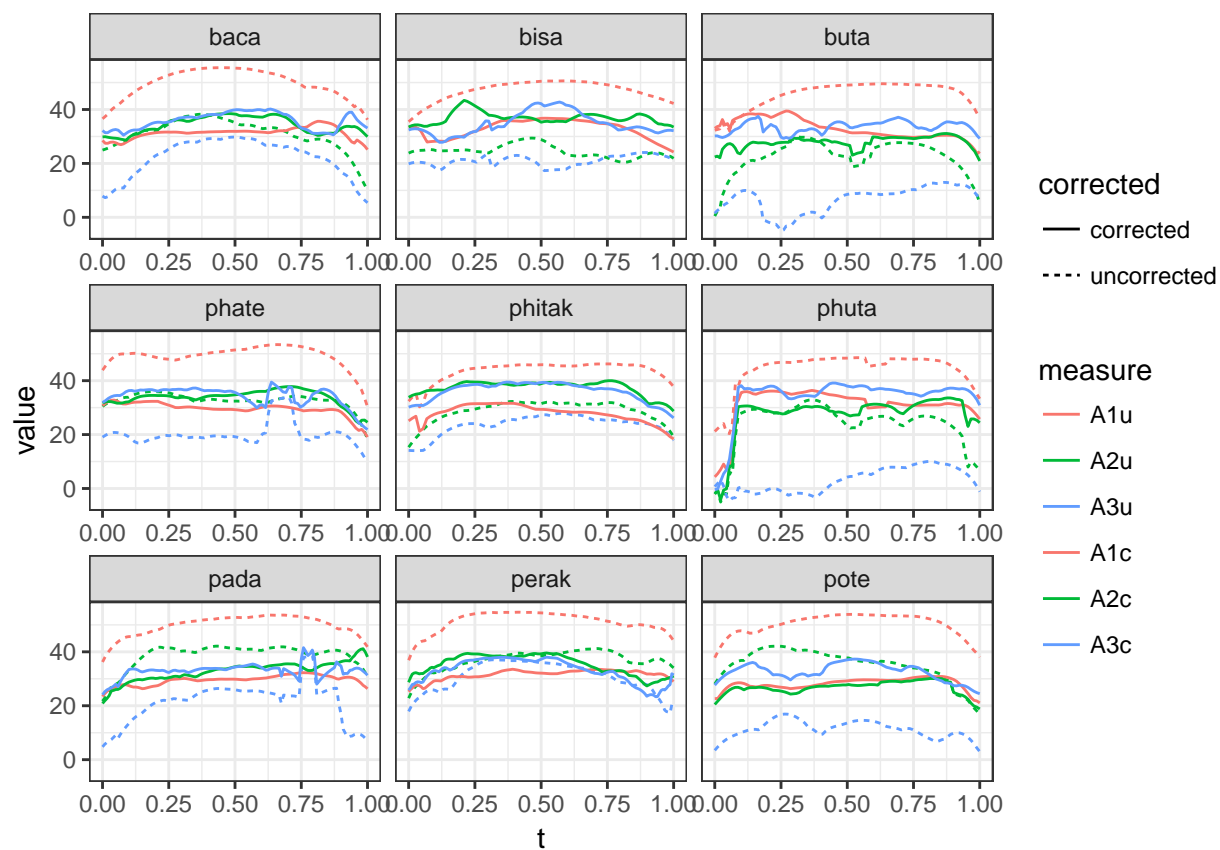


Largest differences for A1 and A2.

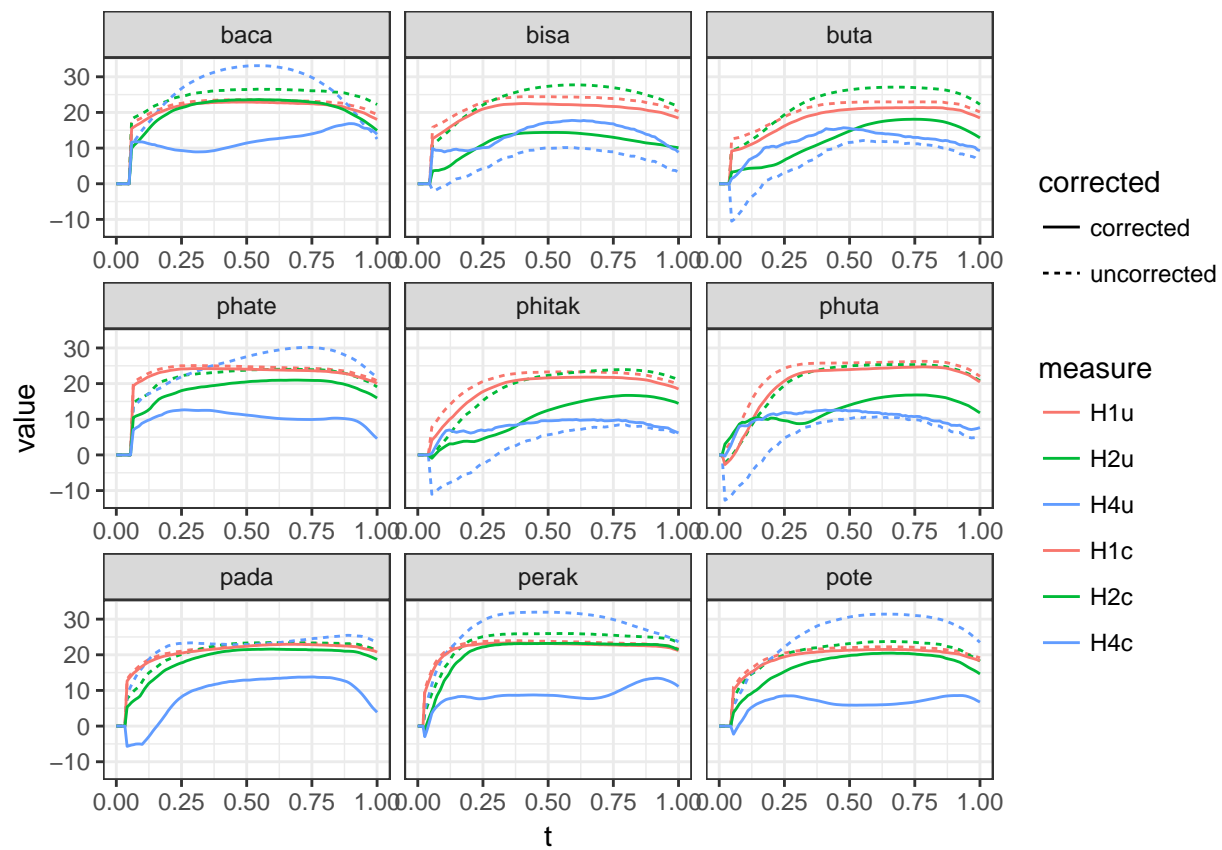
PraatSauce corrected vs. uncorrected, H1/H2/H4, estimated bandwidths



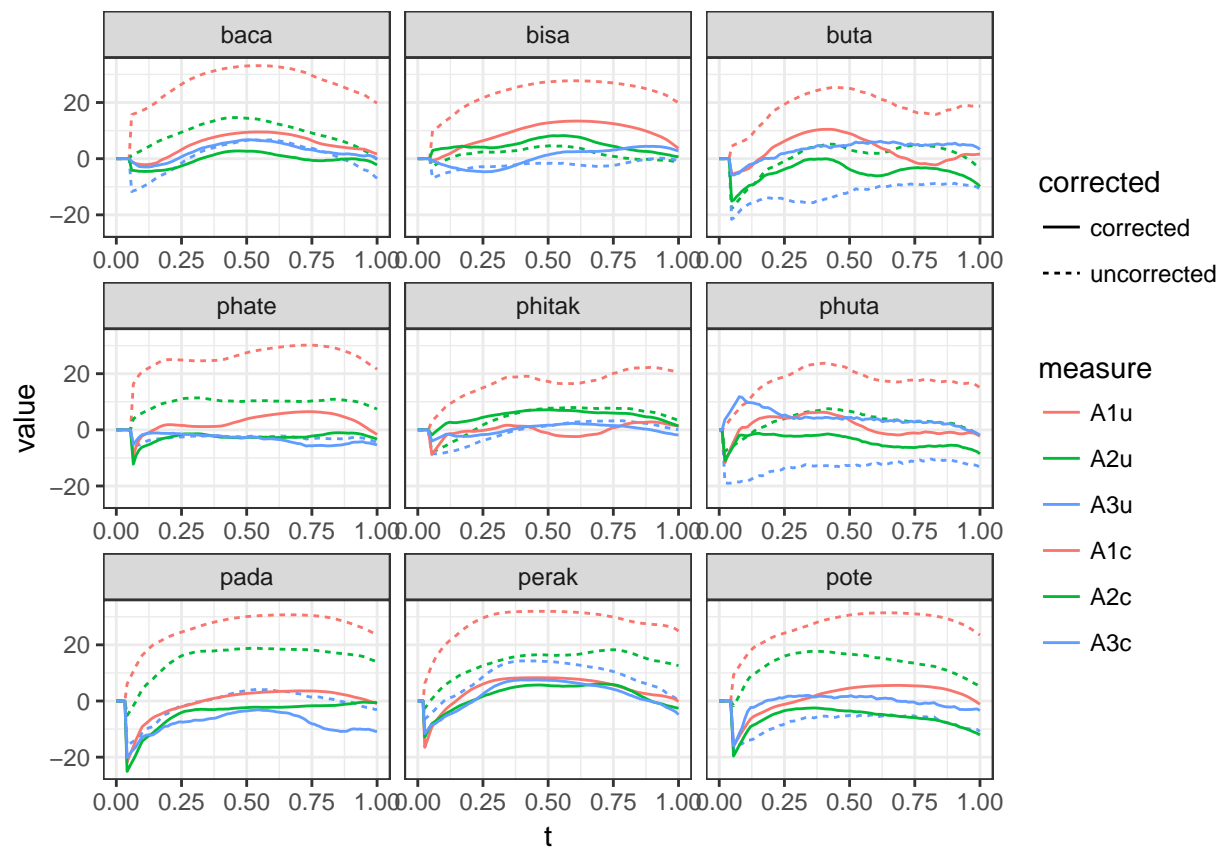
PraatSauce corrected vs. uncorrected, A1/A2/A3, estimated bandwidths



VoiceSauce corrected vs. uncorrected, H1/H2/H4, formula bandwidths



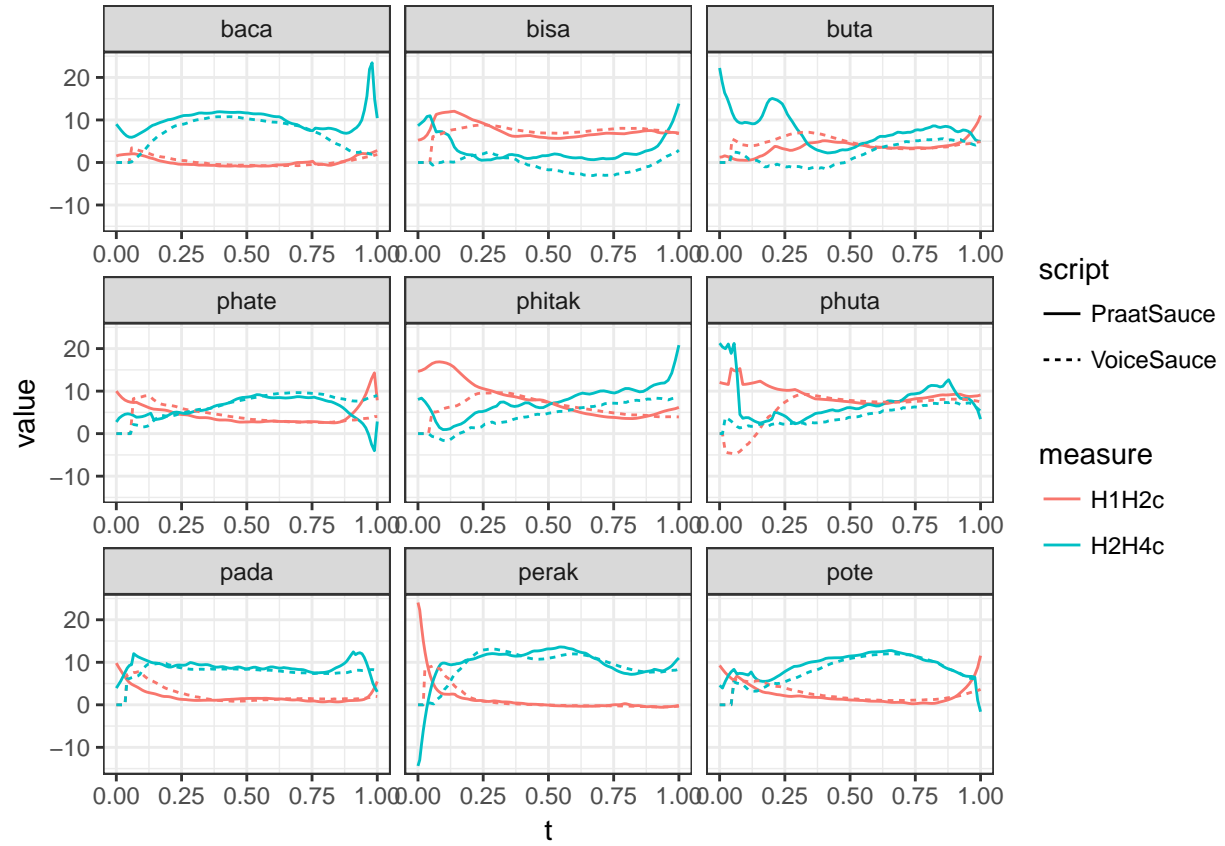
VoiceSauce corrected vs. uncorrected, A1/A2/A3, formula bandwidths



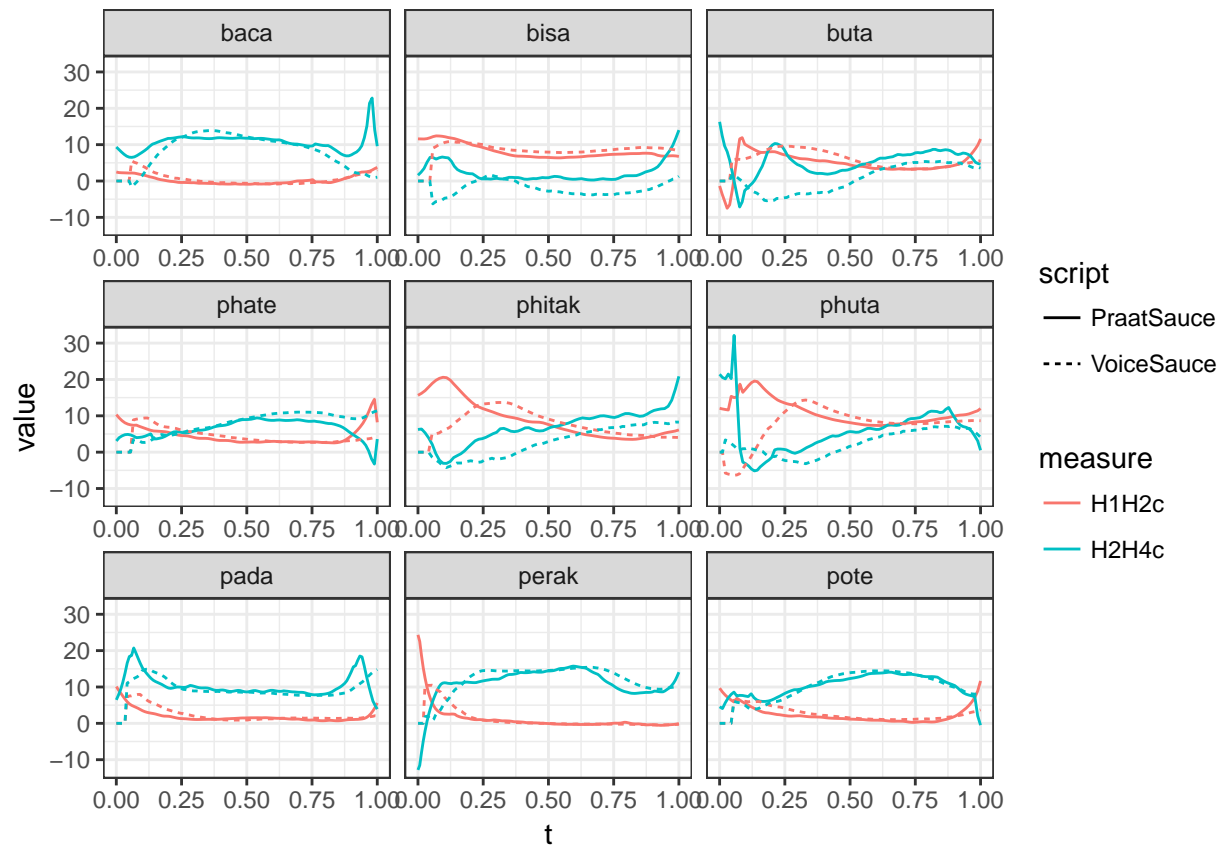
Corrected differences

More interesting is probably a comparison of the corrected differences:

PraatSauce vs. VoiceSauce estimated bandwidths, H1*-H2* & H2*-H4*

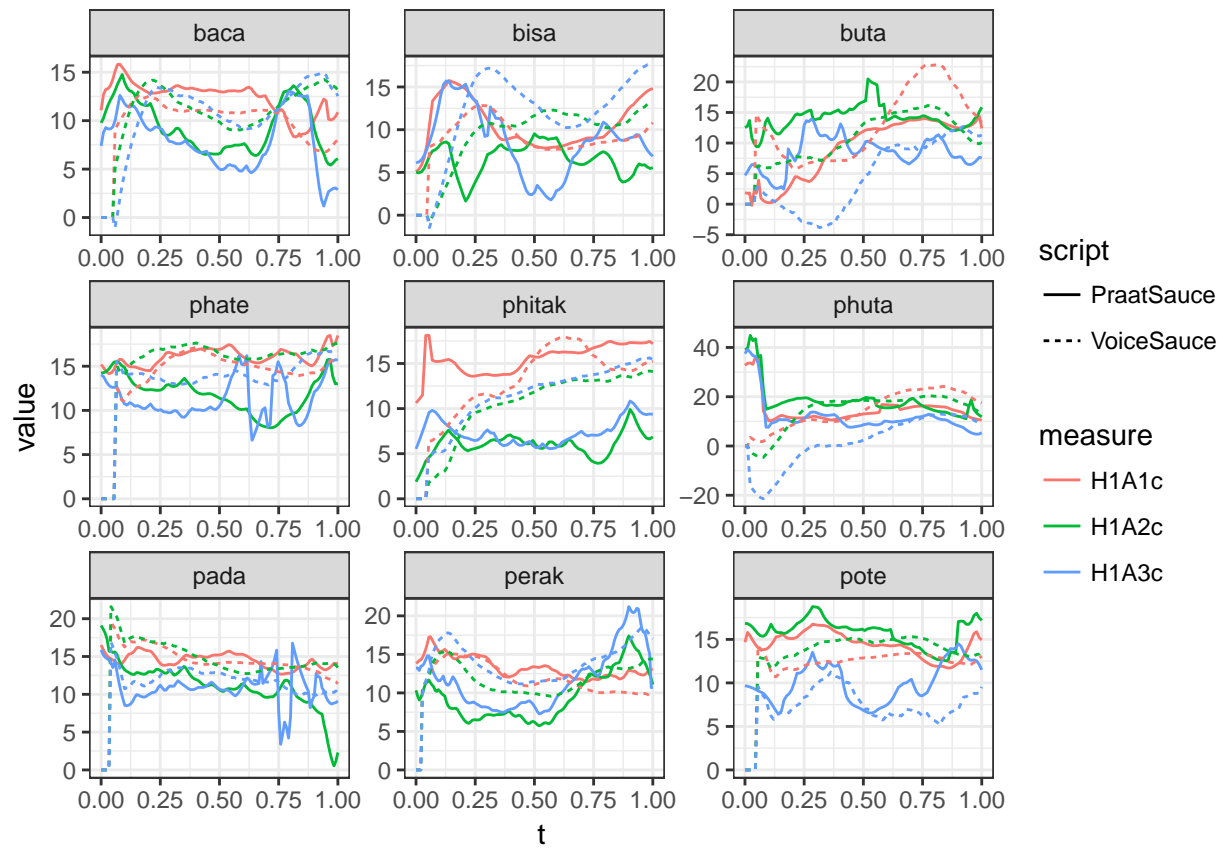


PraatSauce vs. VoiceSauce formula bandwidths, H1*-H2* & H2*-H4*

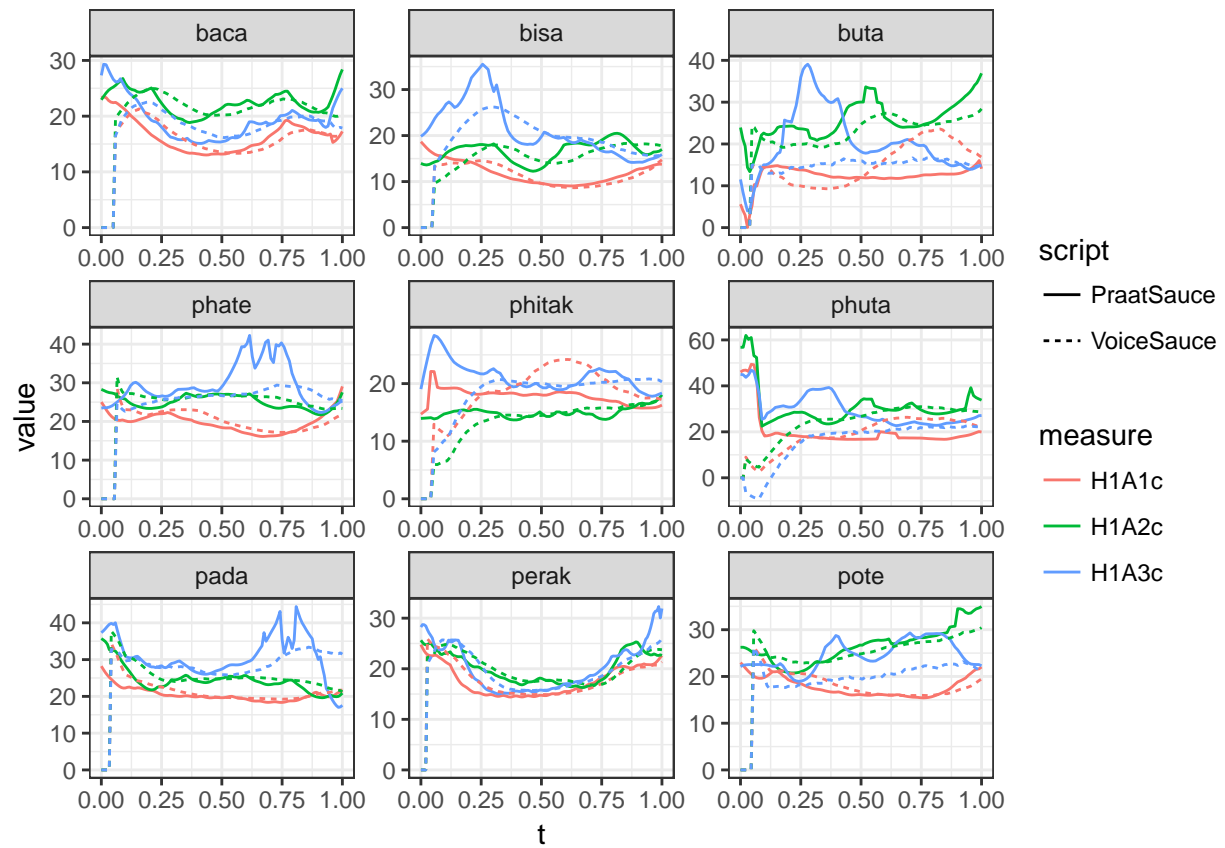


Very similar.

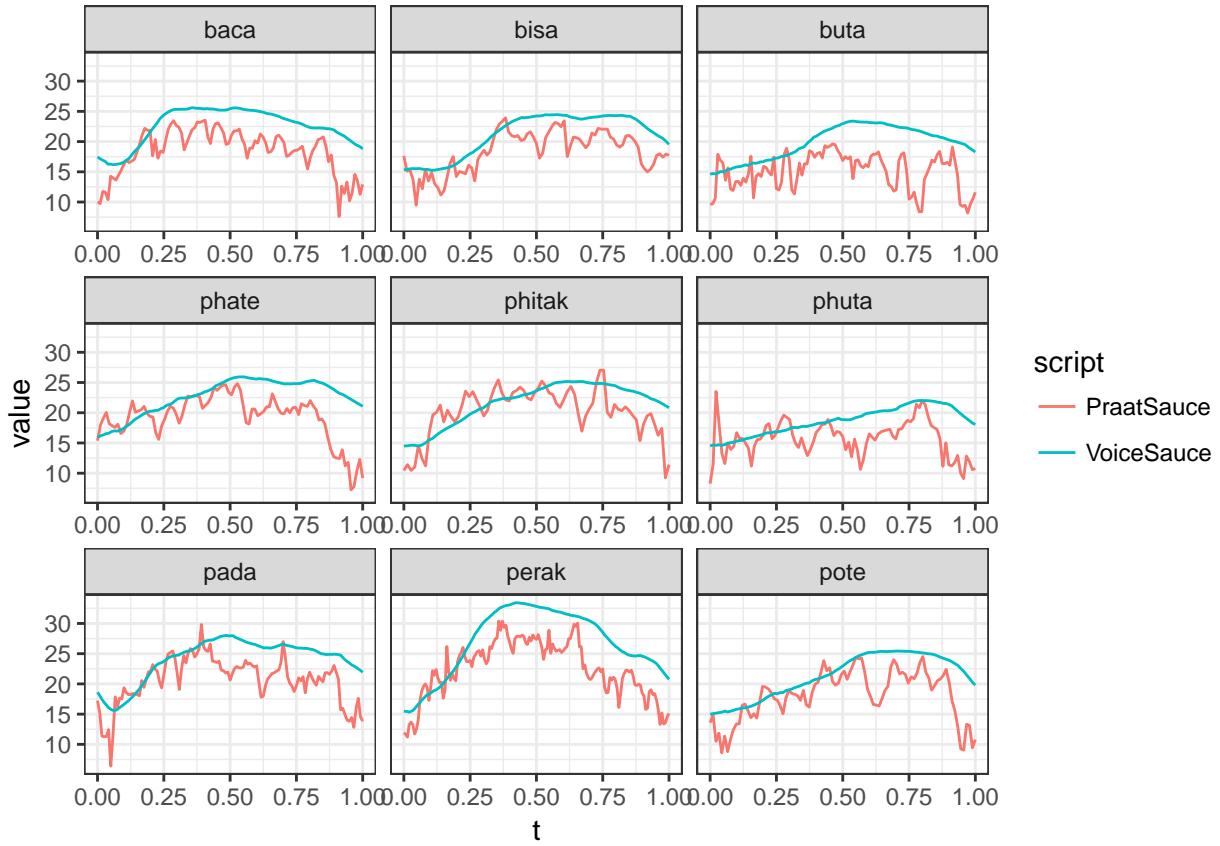
PraatSauce vs. VoiceSauce: estimated bandwidths, H1*-A1*, A2*, A3*



PraatSauce vs. VoiceSauce: formula bandwidths, H1*-A1*, A2*, A3*



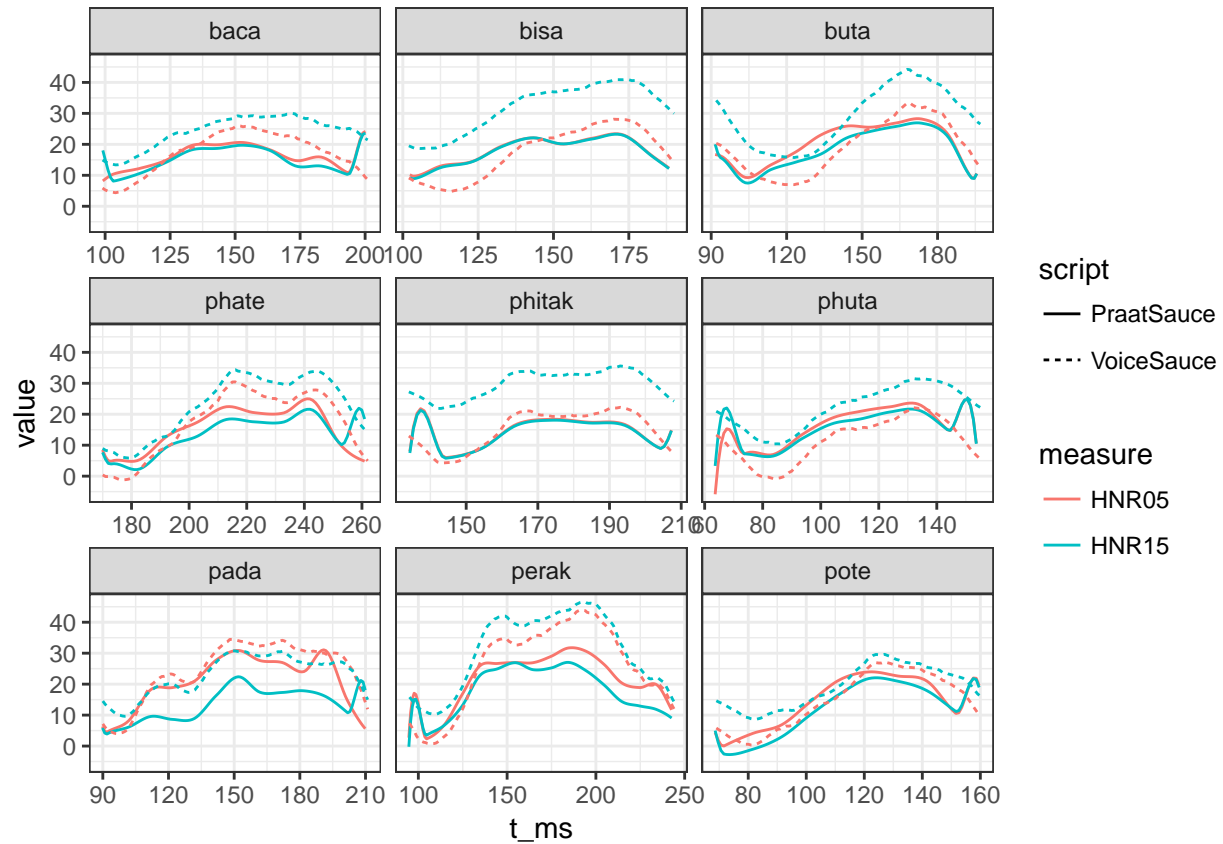
Cepstral peak prominence



Praat(Sauce) estimates are comparable if smoothed.

Harmonic to noise ratios

Here just showing HNR05 and HNR15 for clarity.

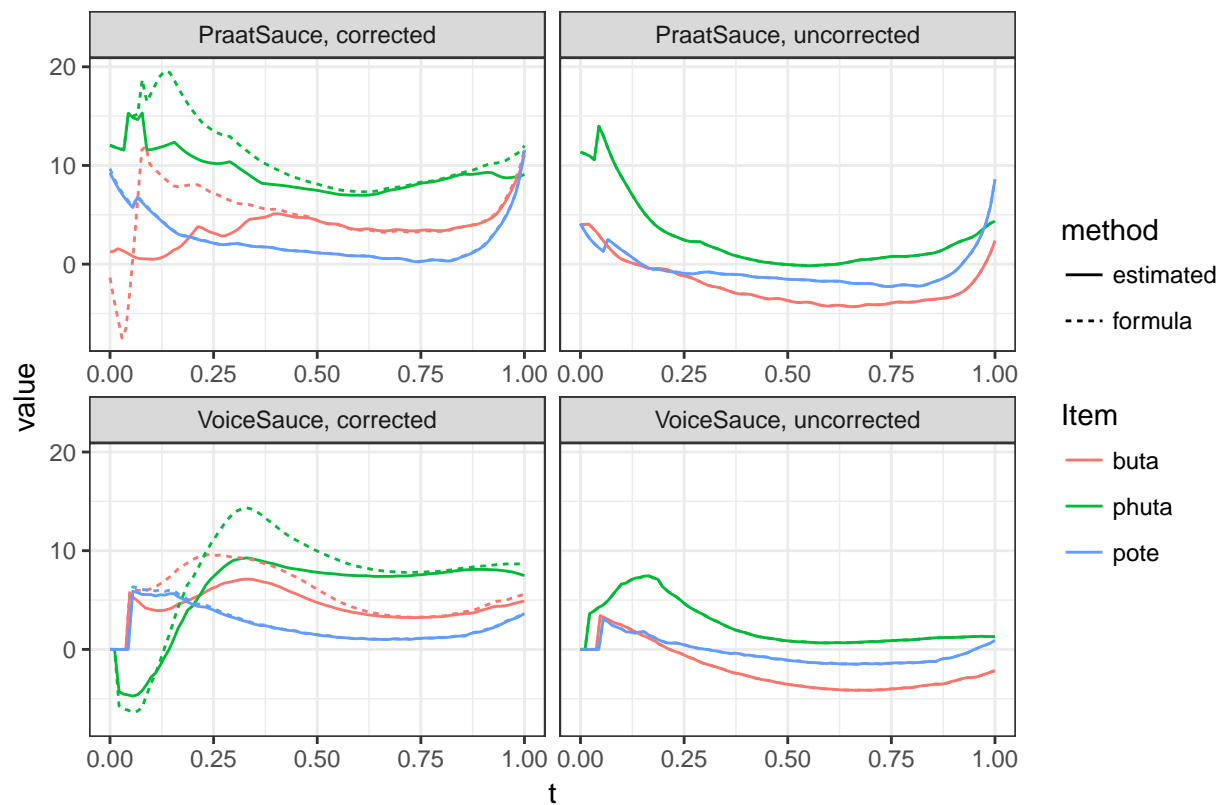


Again, the Praat estimates differ in amplitude, but maintain roughly the same trajectories. However, the PraatSauce implementation is much less sophisticated than that of VoiceSauce, and relies entirely on Praat's `To Harmonicity...` function. There does not appear to be much difference in the bands as PraatSauce estimates them.

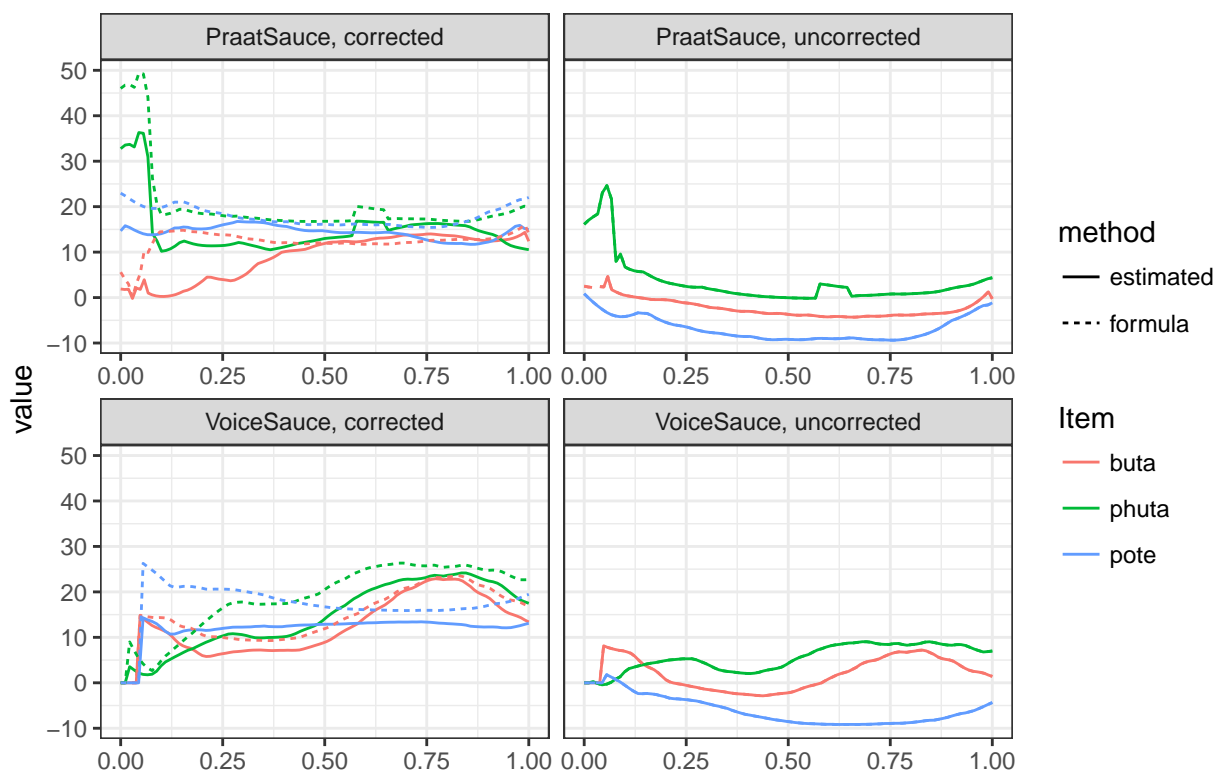
Distinguishing vowel qualities

High/back vowels

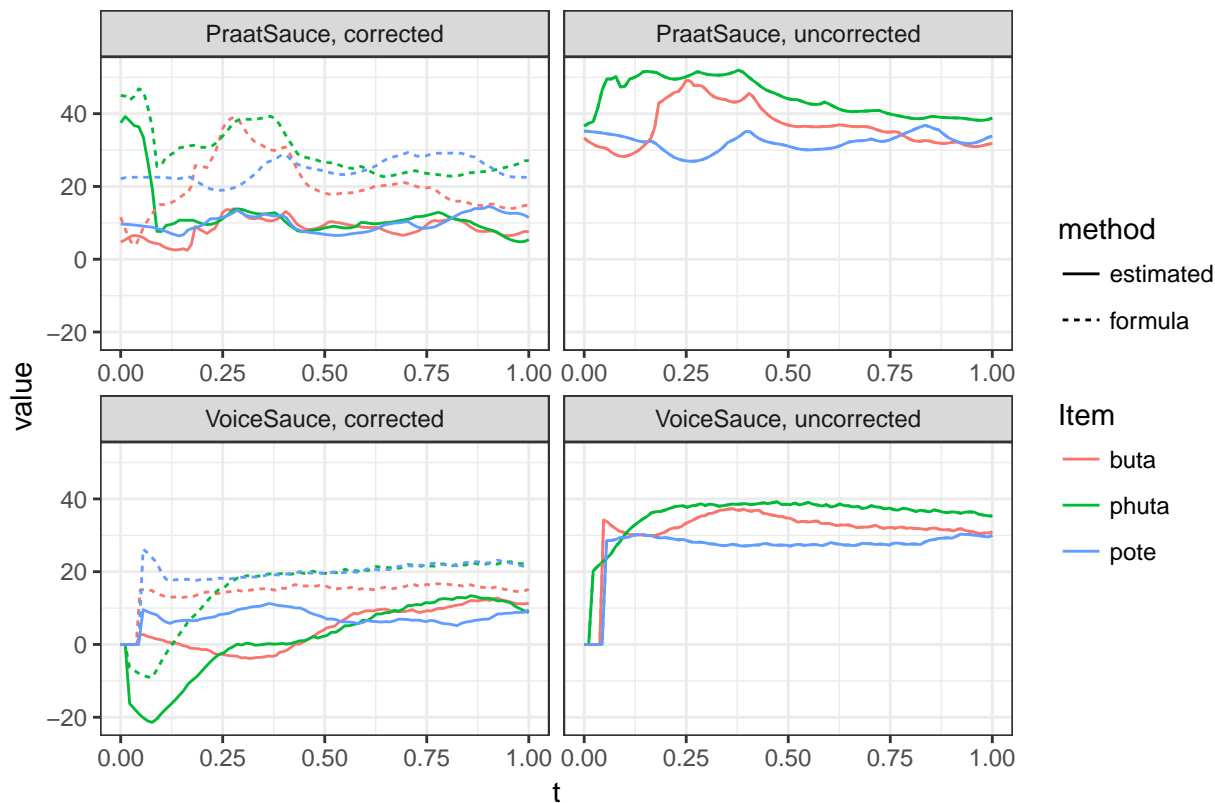
VoiceSauce vs. PraatSauce H1*–H2*



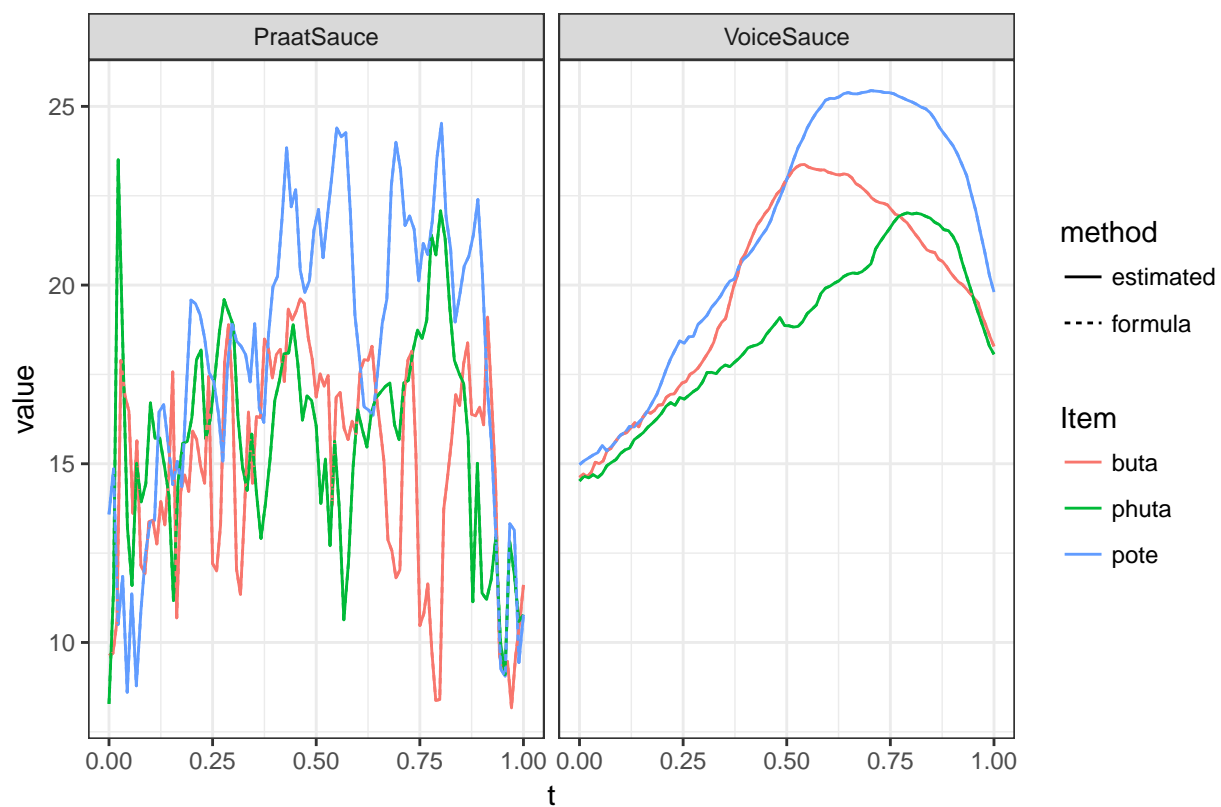
VoiceSauce vs. PraatSauce H1*-A1*



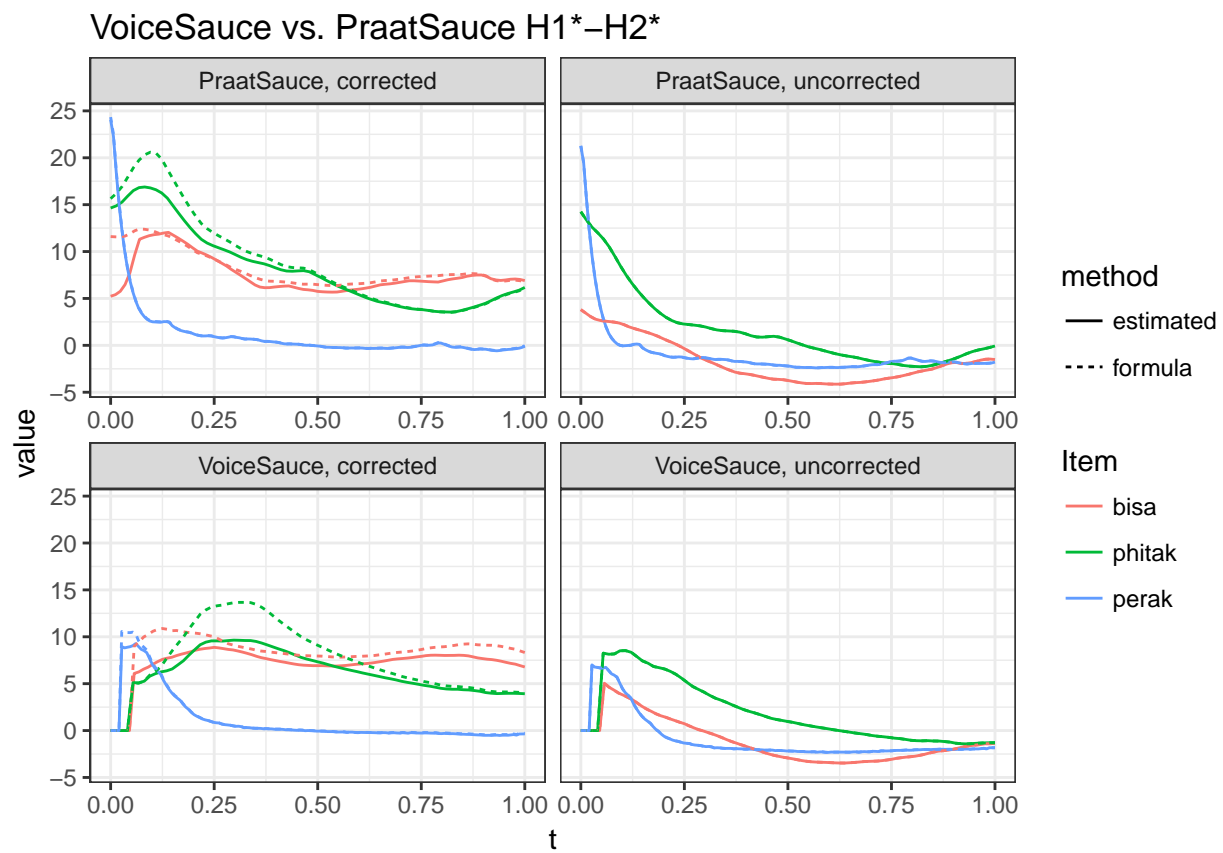
VoiceSauce vs. PraatSauce H1*-A3*



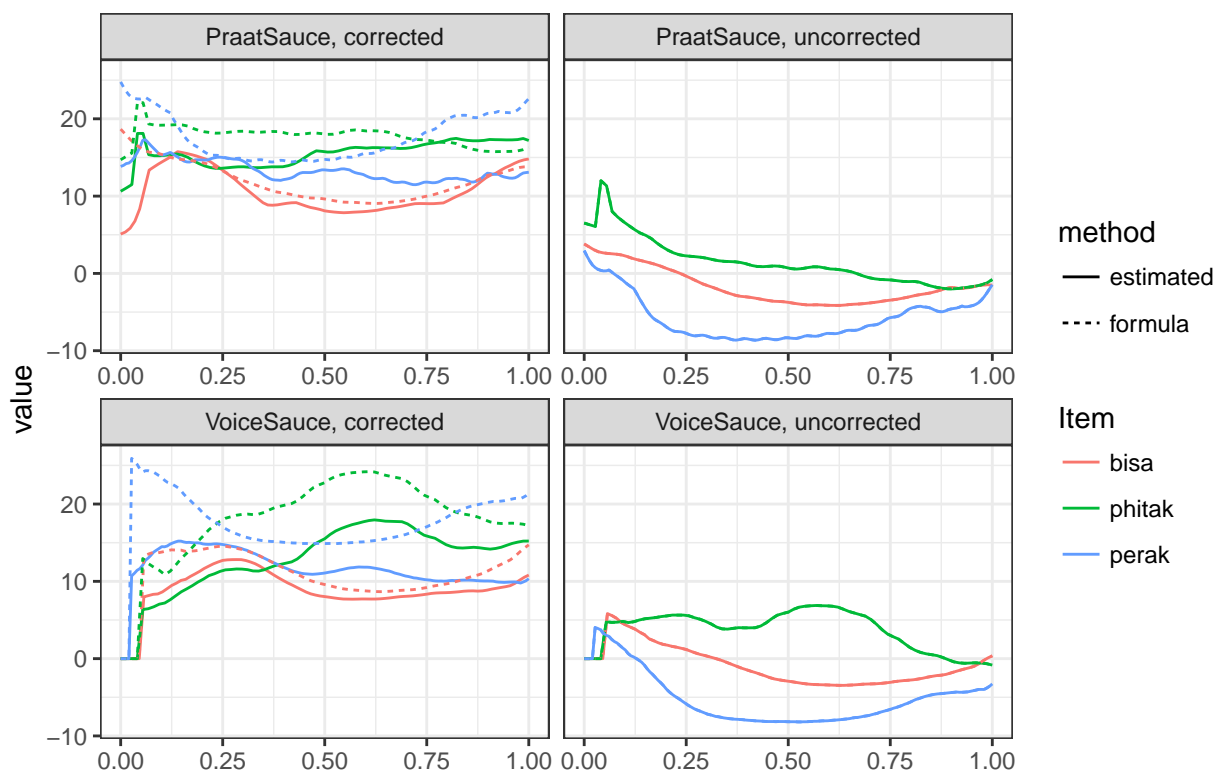
VoiceSauce vs. PraatSauce CPP



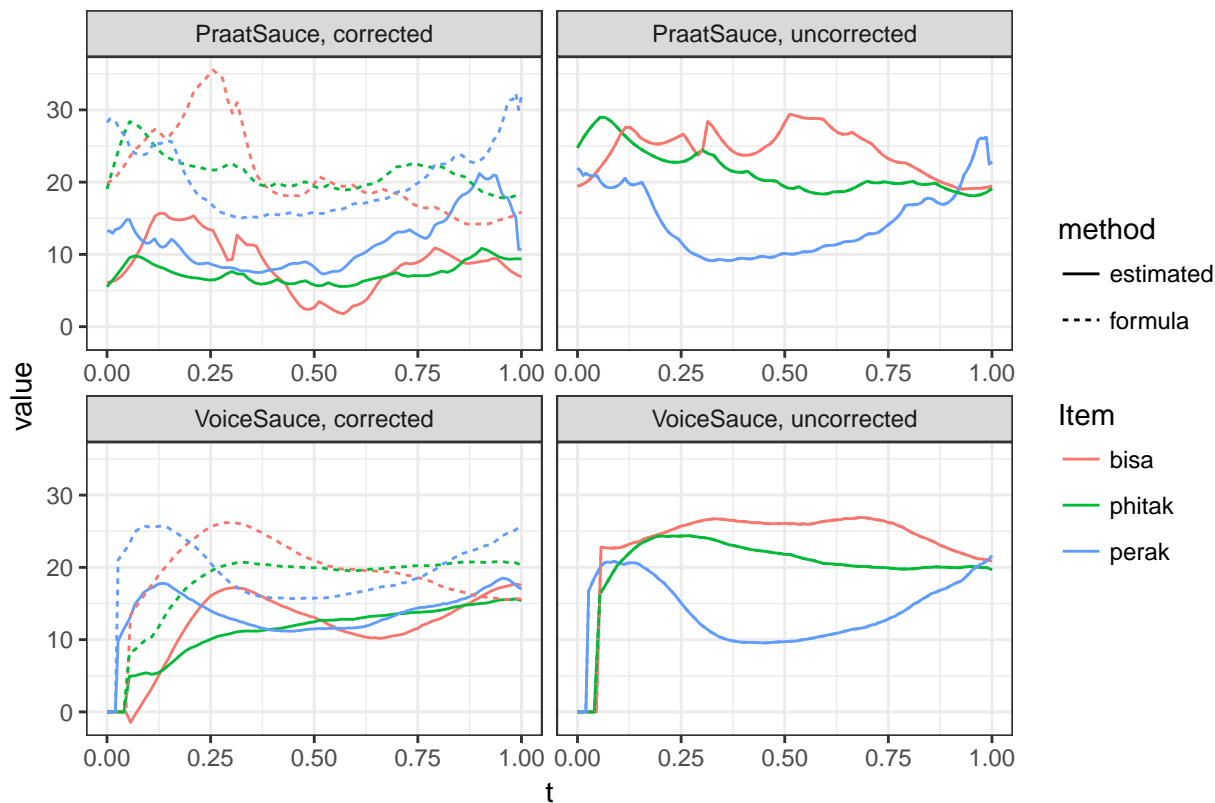
High/front vowels



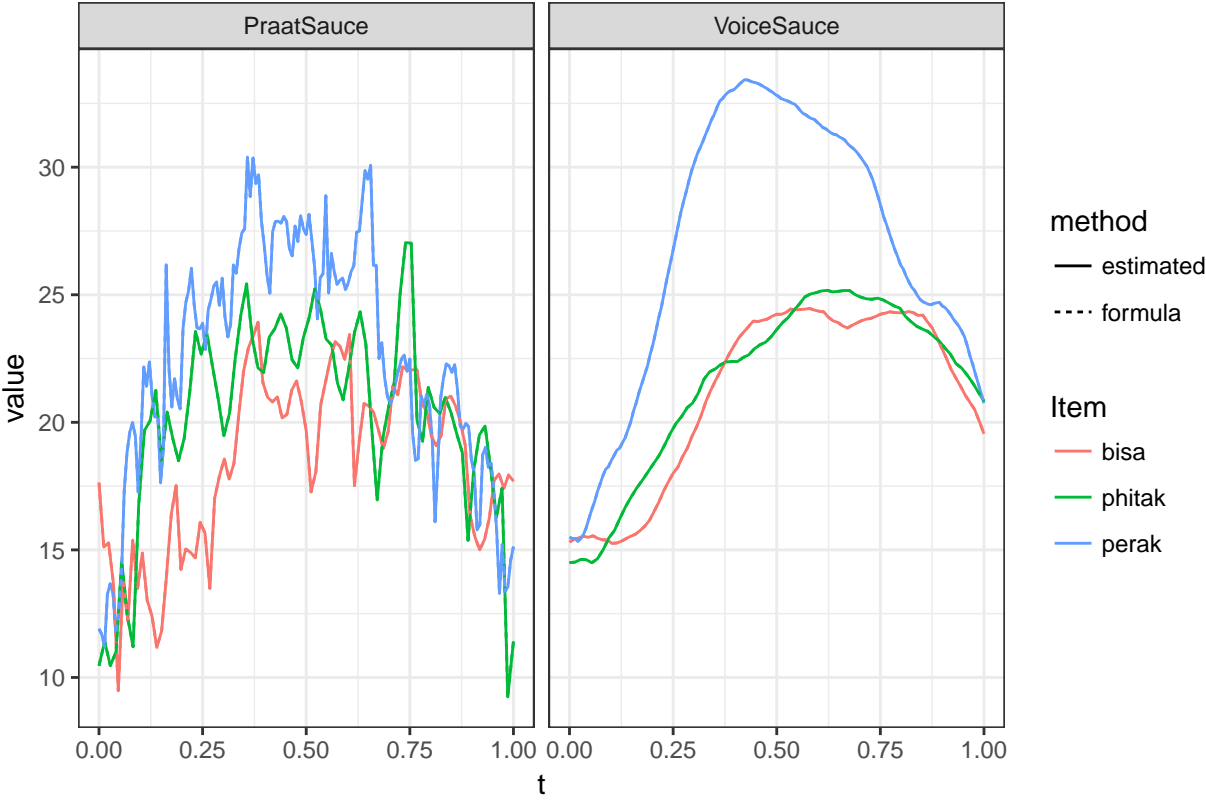
VoiceSauce vs. PraatSauce H1*–A1*



VoiceSauce vs. PraatSauce H1*–A3*

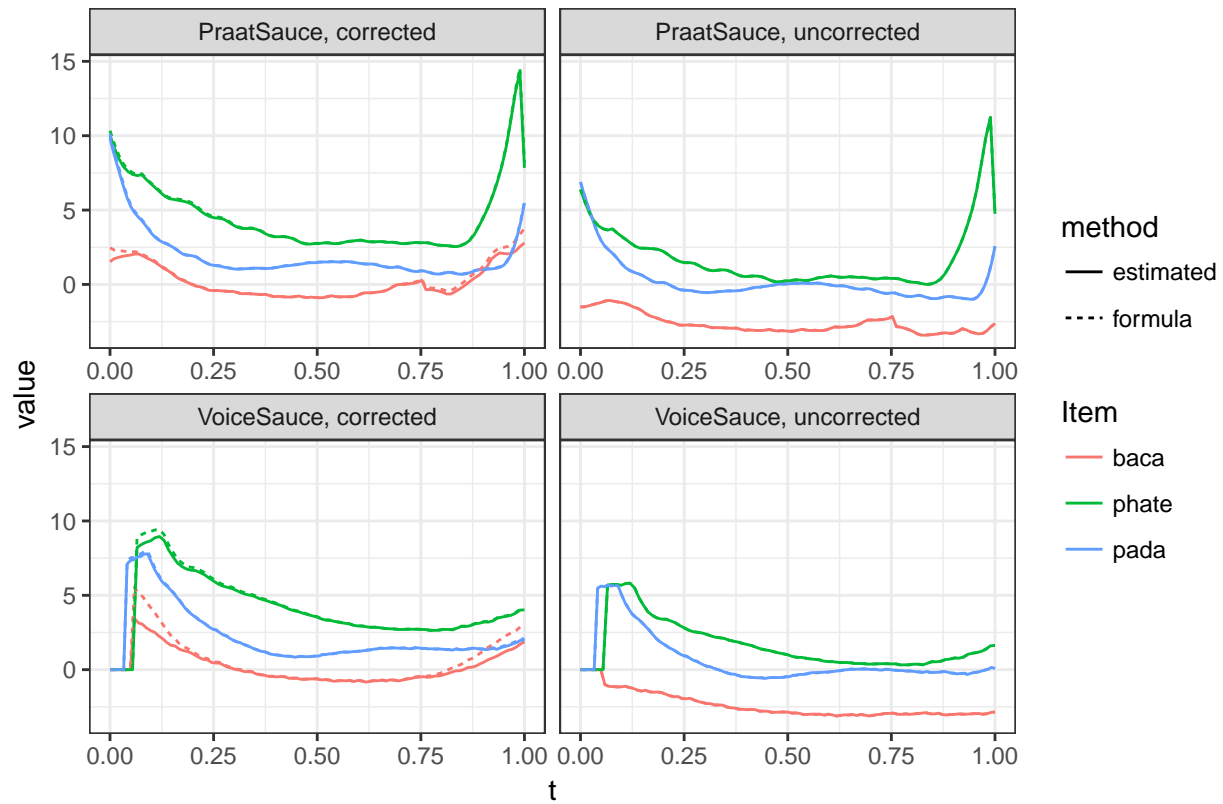


VoiceSauce vs. PraatSauce CPP

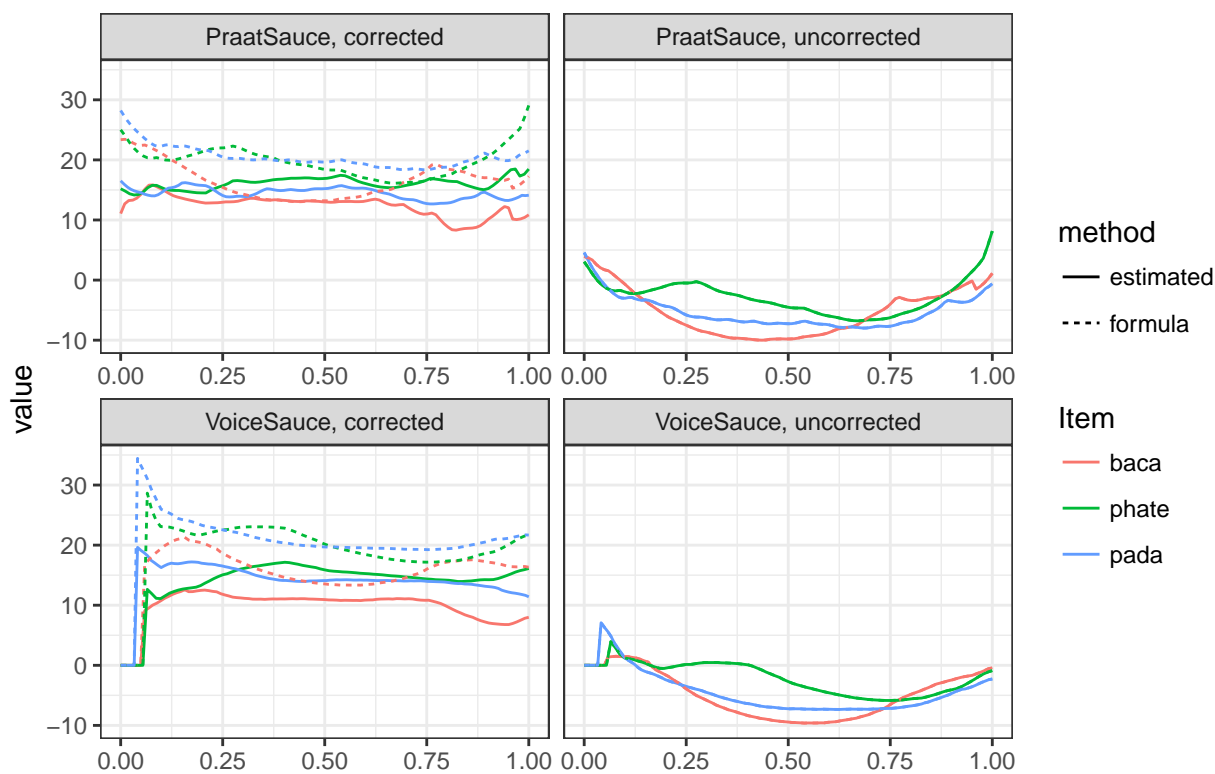


Low/front vowels

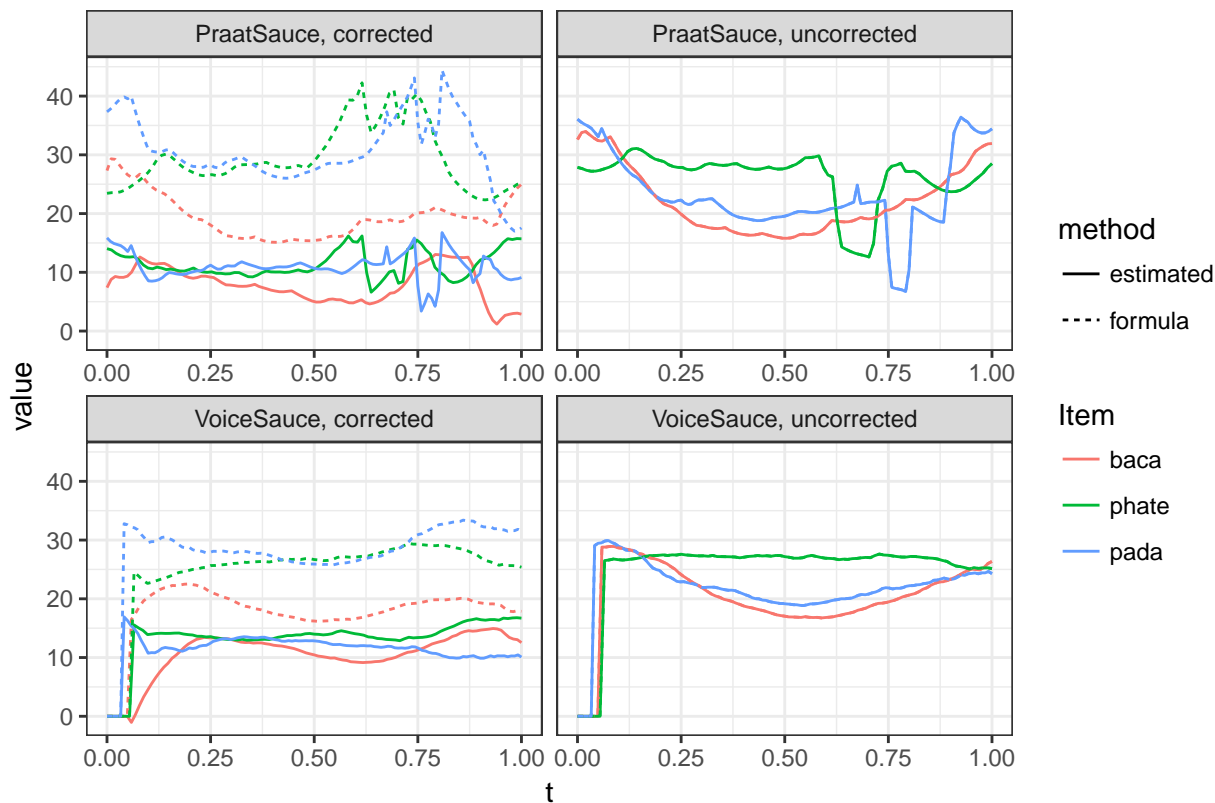
VoiceSauce vs. PraatSauce H1*–H2*



VoiceSauce vs. PraatSauce H1*-A1*



VoiceSauce vs. PraatSauce H1*-A3*



VoiceSauce vs. PraatSauce CPP

