The University of Manchester

Quantifying vocal fold activity: two new methods for analysing electroglottographic data

Stefano Coretta

The University of Manchester

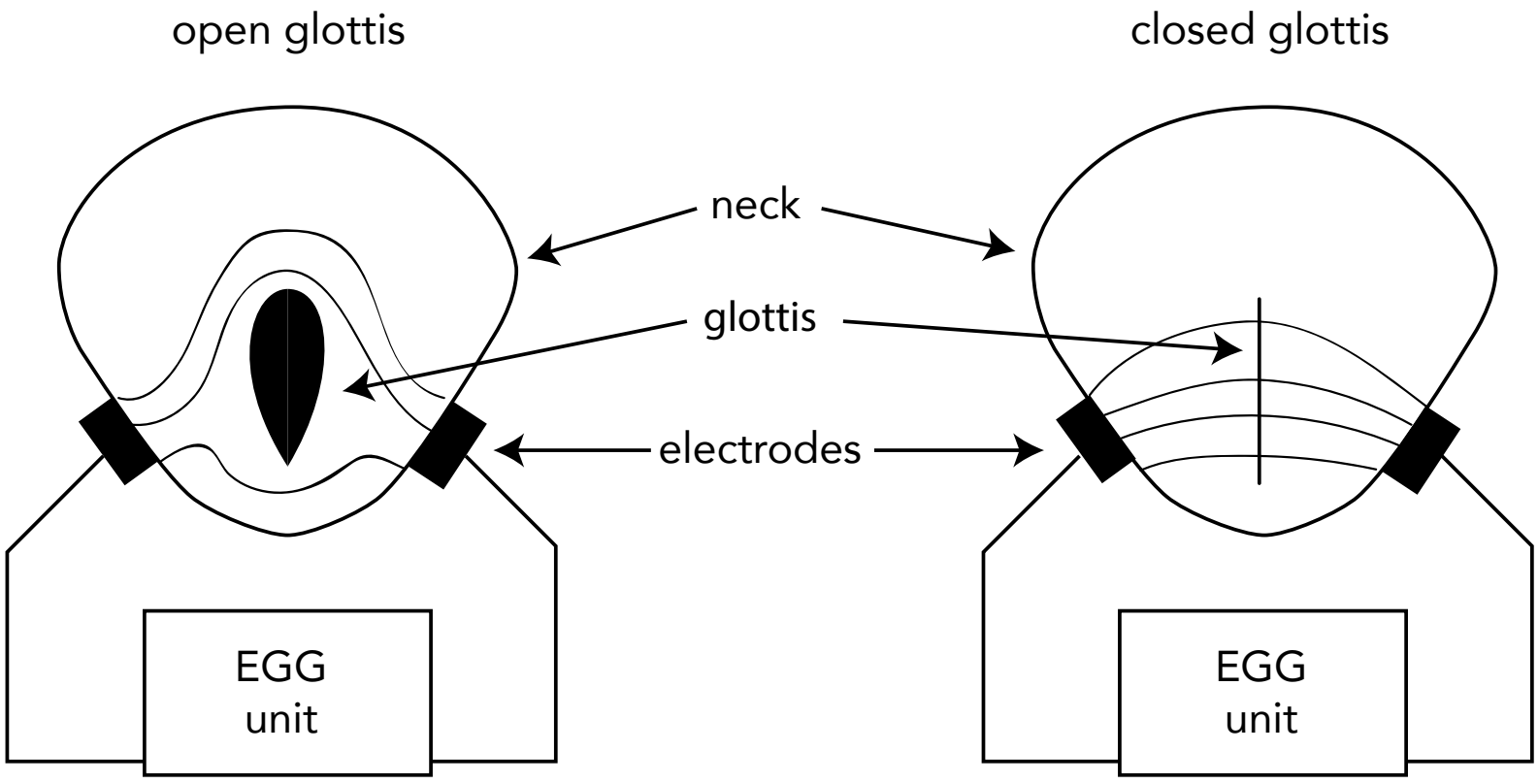
New developments in speech sensing and imaging, Lisbon, 23rd June 2018

Background: The technique

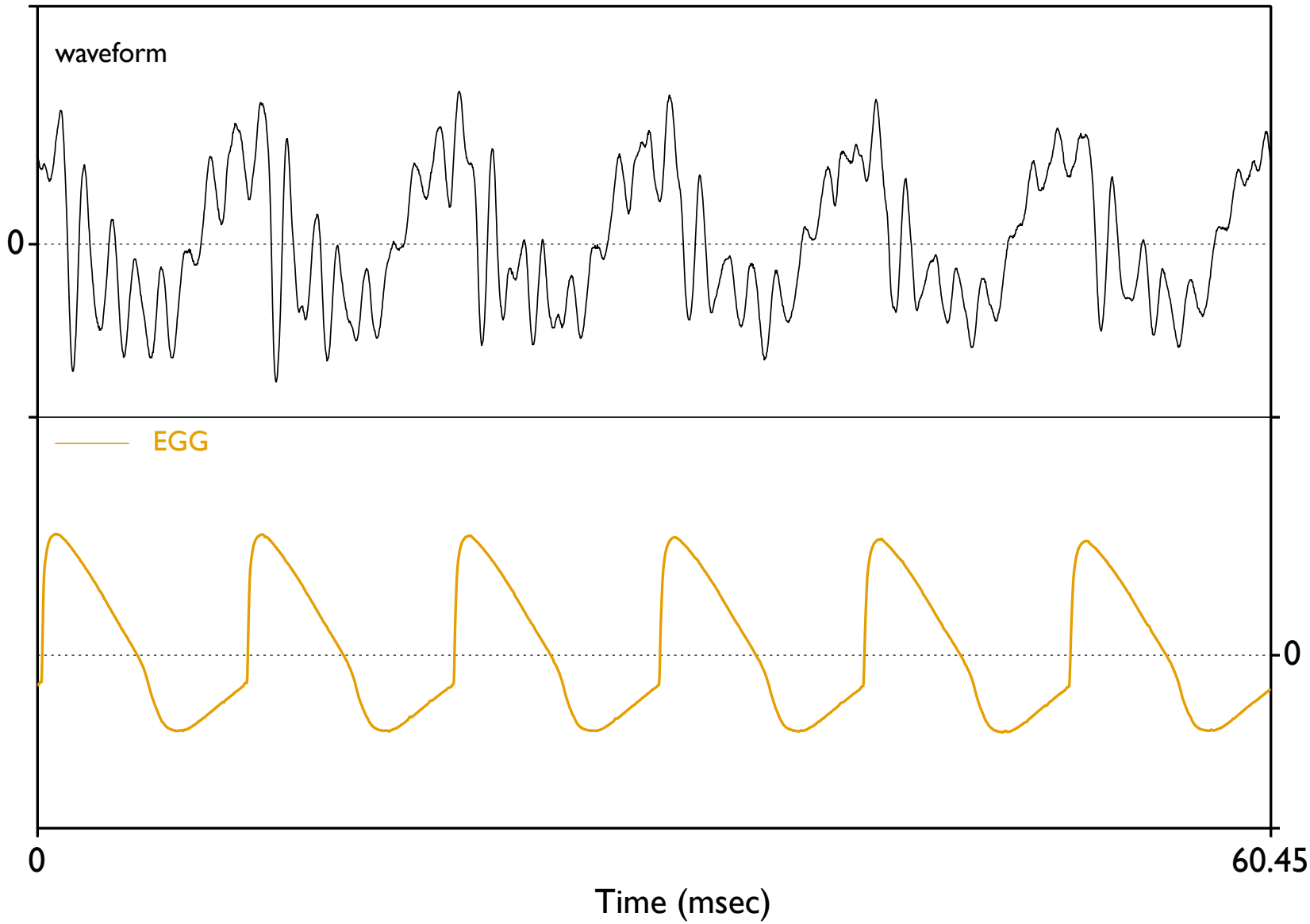
- EGG (Fabre, 1957; Scherer & Titze, 1987; Rothenberg & Mahshie, 1988)
- **Purpose**: estimation of vocal folds contact area (VFCA)
- **How**: based on modulations of a current that travels the neck generated by the opening and closing of the vocal folds

Background: The technique

Background: The technique



Background: The technique



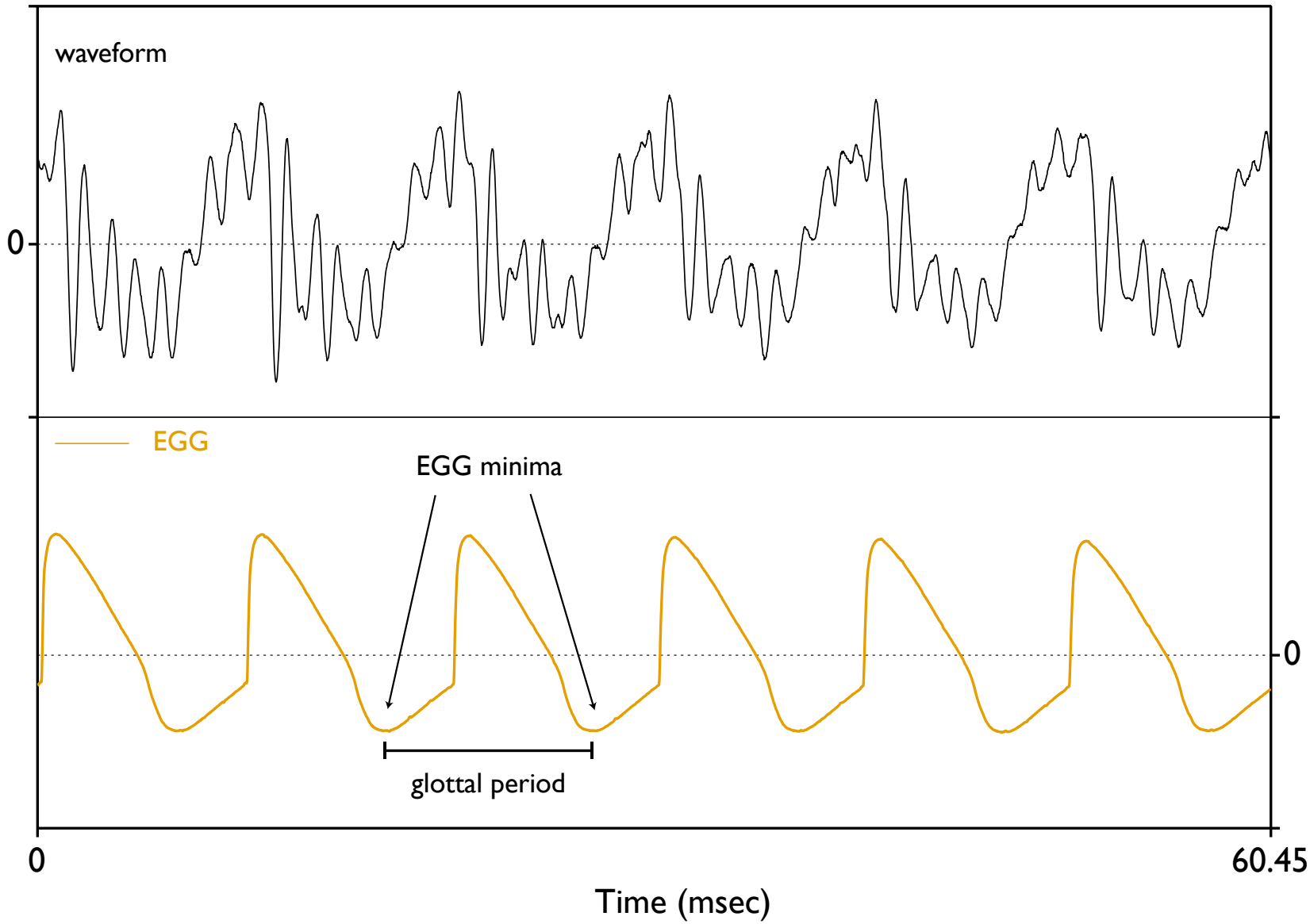
Background: The technique

- Pros:
  - non-invasive
  - relatively simple signal
- Cons:
  - approximantion of VFCA (Herbst et al., 2014; Hampala et al., 2016)
- Use:
  - estimation of vocal fold activity
  - estimation of fundamental frequency
  - study of pathological speech

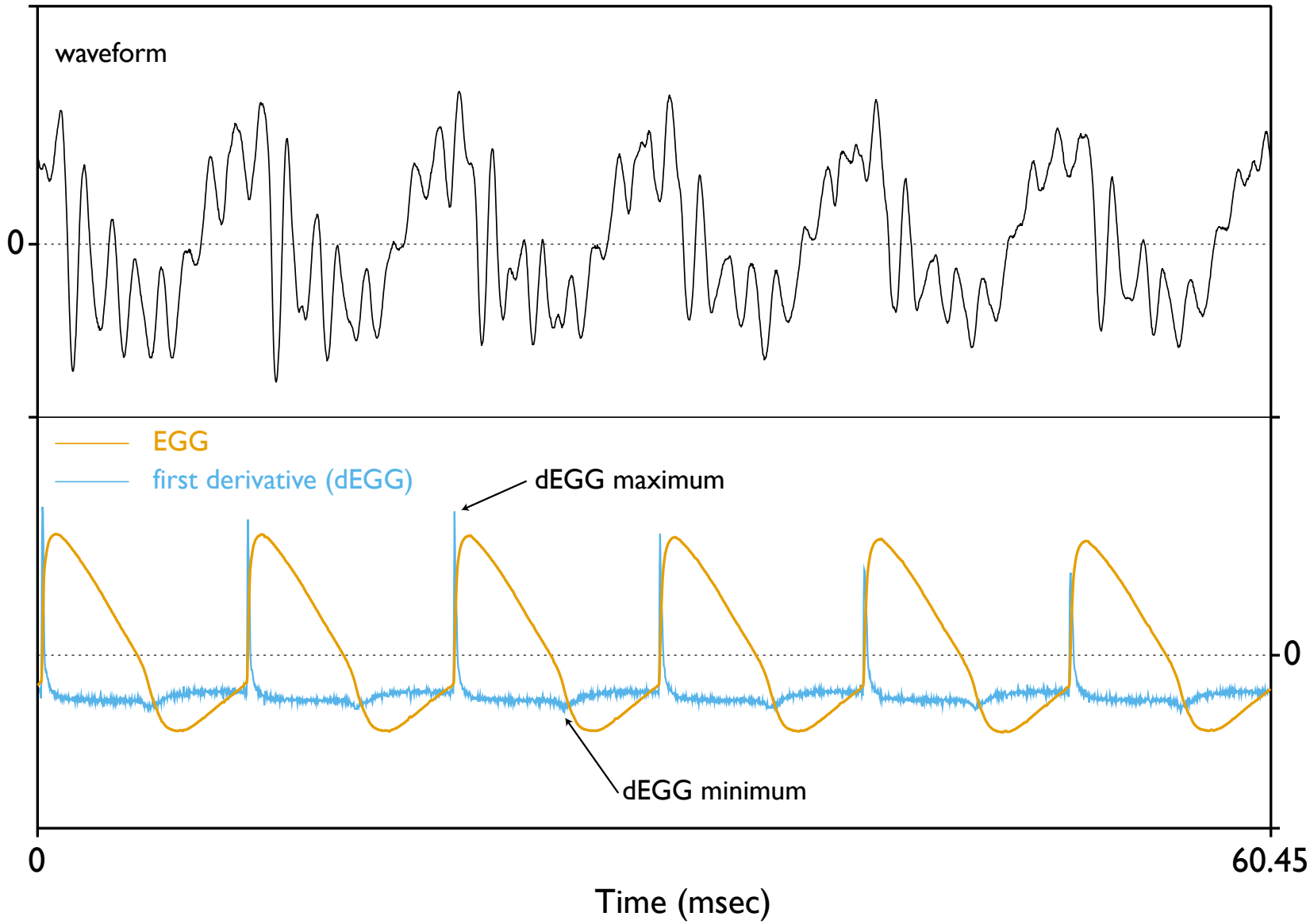
Background: The signal

- **Contact Quotient** (Awan et al., 2015; Herbst et al., 2017)
  - proportion of the contact phase relative to the glottal period
- **Wavegram** (Herbst et al., 2010)
  - visualisation of amplitude changes in the signal through time

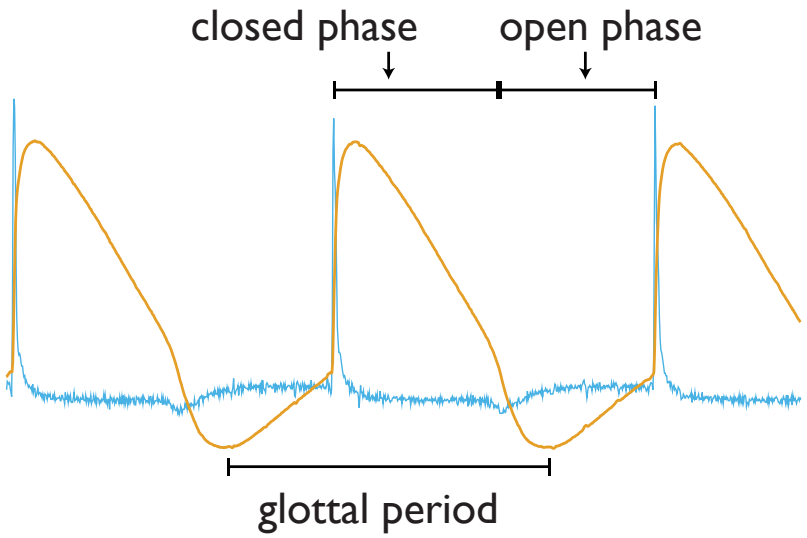
Background: The signal



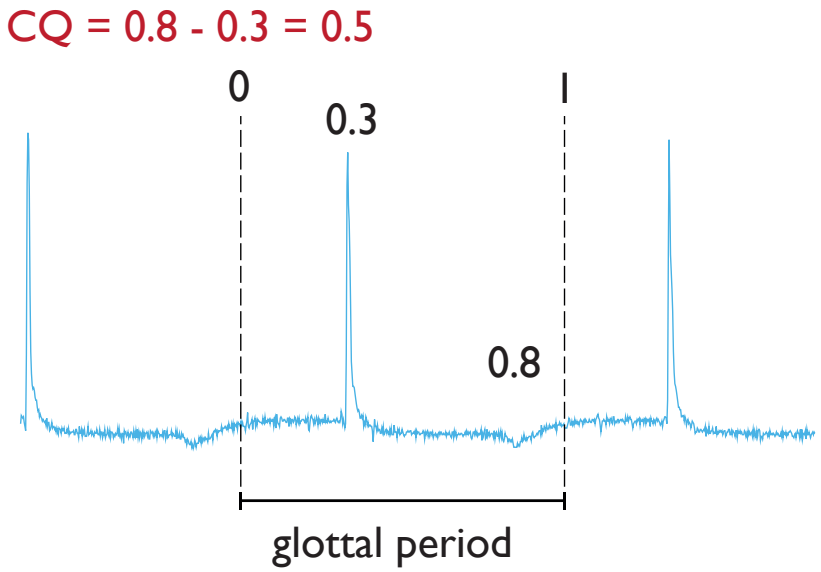
Background: The signal



Background: The signal



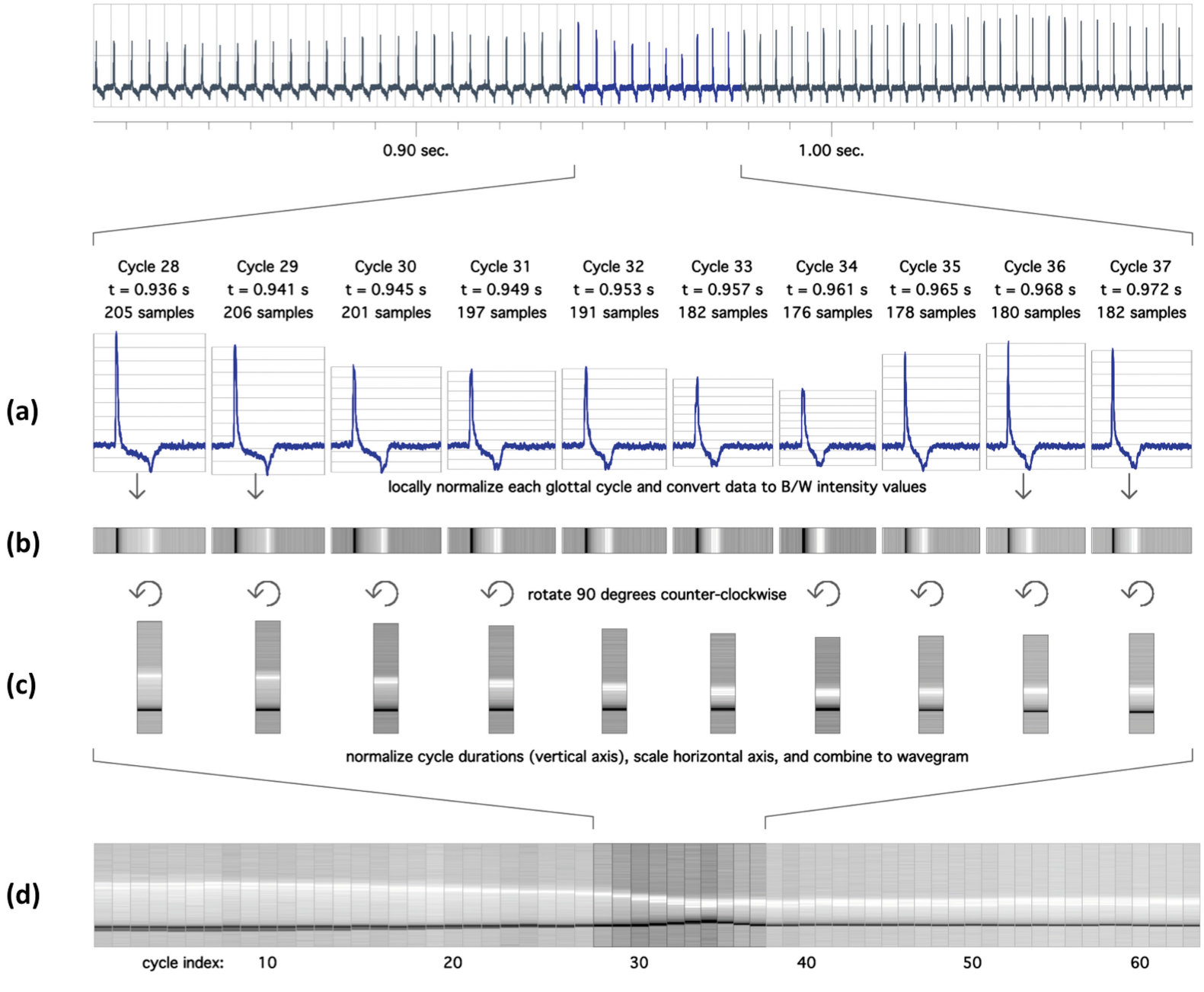
Background: The signal



Background: The signal

- CQ reduces dimentionality of EGG signal
- Herbst et al. (2010) propose the **wavegram** as a multidimentional account of the EGG signal

Background: The signal



Background: This study

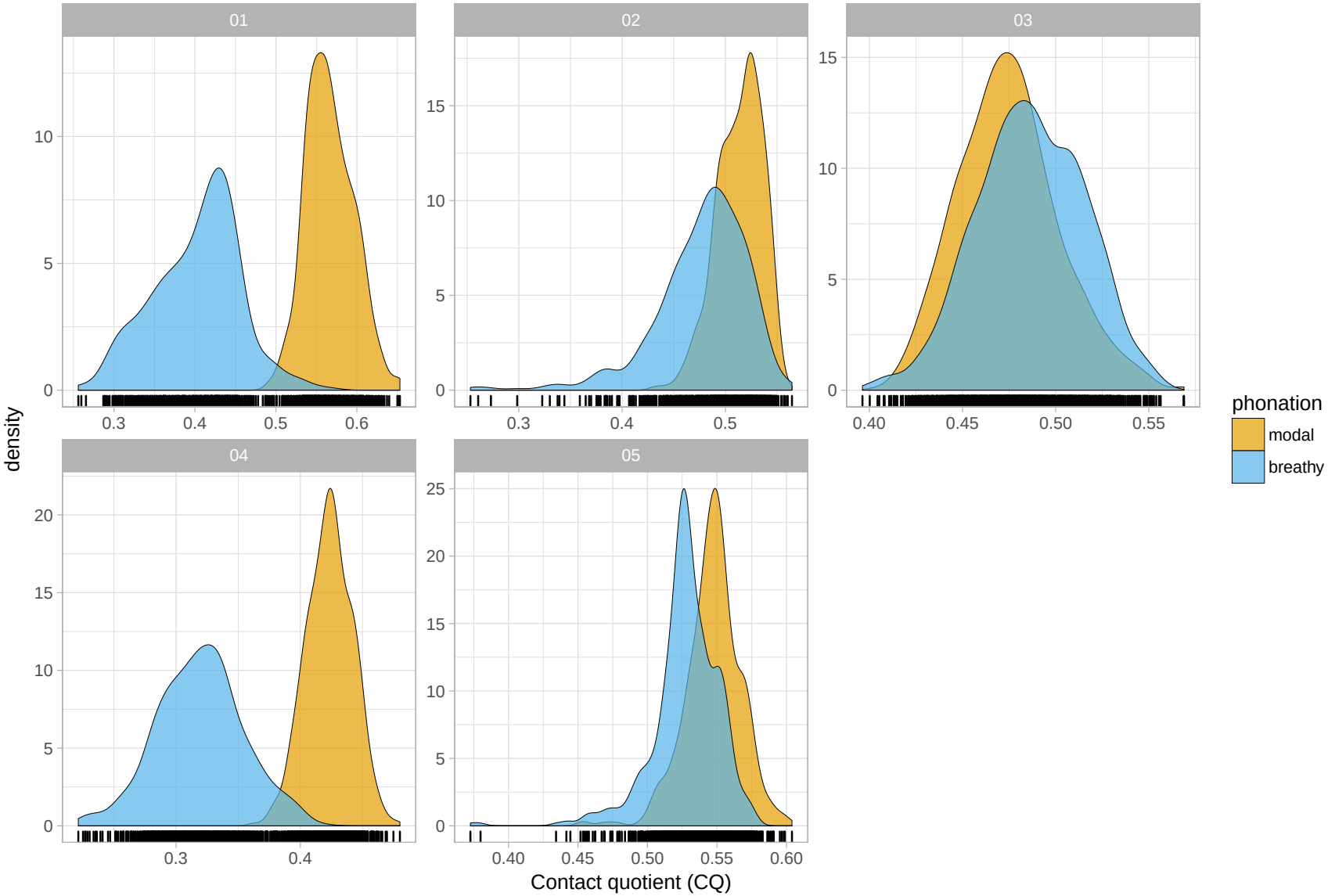
- Assessment of previous methods
  - CQ is not precise (Baken, 1992; Herbst et al., 2017)
  - wavegram cannot be assessed statistically
- Two new techniques
  - wavegram GAMs
  - tracegram

Methods

- 5 phonetically trained speakers (1 F, 4 M, languages: BE, IT)
- [ɑ]/[a] in modal and breathy voice
  - 10 × 2 = 20 tokens per speaker
  - 100 tokens
- equipment
  - Glottal Enterprises EG2-PCX2 electroglottograph
  - Movo LV4-O2 Lavalier microphone (sample rate 44100 Hz, 16-bit)
- analysis window
  - 500 ms portion centred around mid point of each token

Results: Contact Quotient (CQ)

Density plots of CQ in modal and breathy phonation by speaker



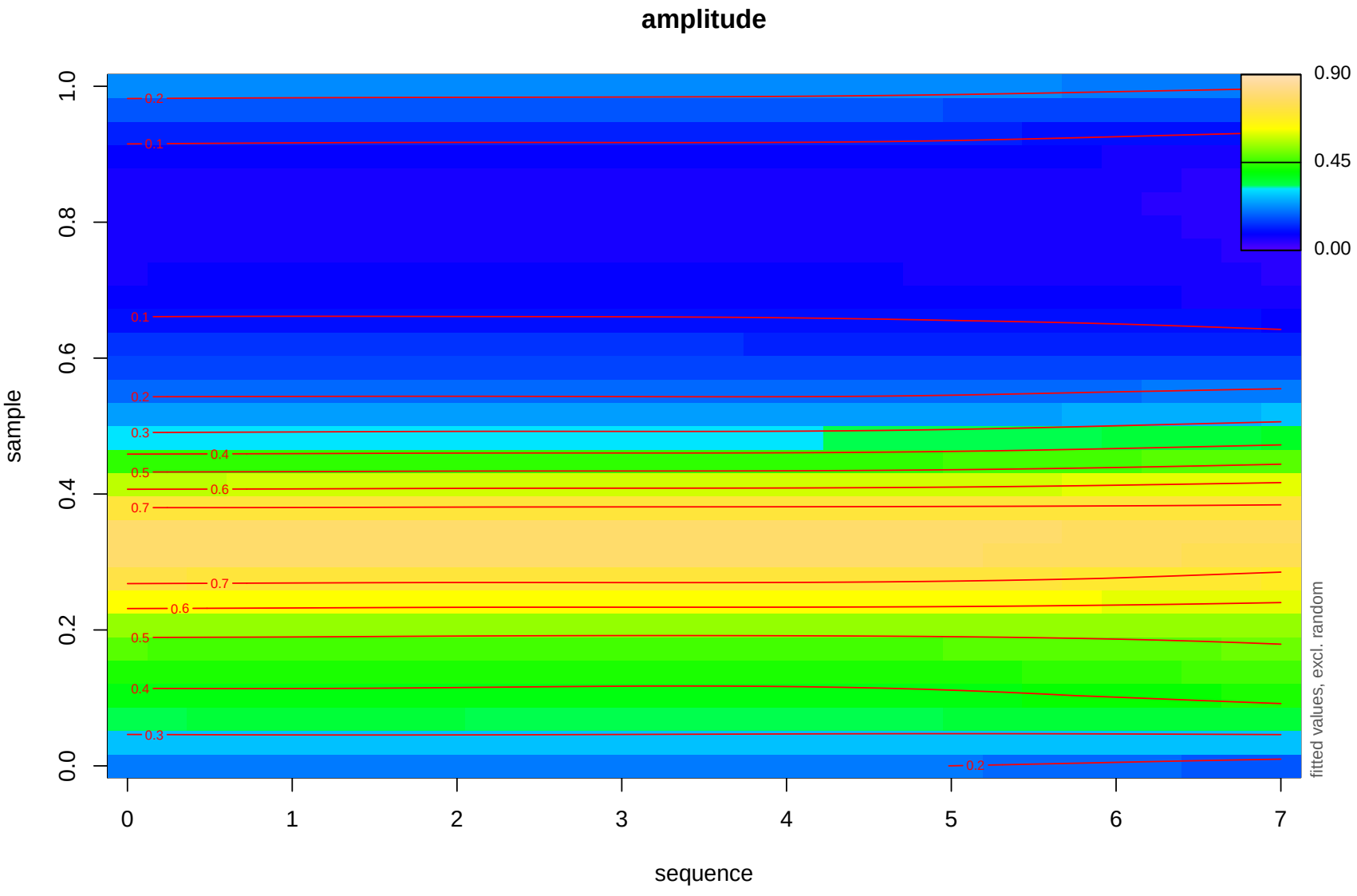
Results: Contact Quotient (CQ)

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: contact_quotient ~ phonation + (1 + phonation | speaker)
##    Data: tracegram
##
## REML criterion at convergence: -24474.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -7.1831 -0.5597  0.0237  0.6202  5.3121
##
## Random effects:
##  Groups   Name             Variance Std.Dev. Corr
##  speaker  (Intercept)      0.003305 0.05749
##           phonationbreathy 0.005009 0.07077  -0.19
##  Residual                  0.000976 0.03124
## Number of obs: 5999, groups:  speaker, 5
##
## Fixed effects:
##                  Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)       0.50512    0.02572  4.00001  19.643 3.96e-05 ***
## phonationbreathy -0.06246    0.03166  3.99996  -1.973     0.12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## phontnbrthy -0.190
```

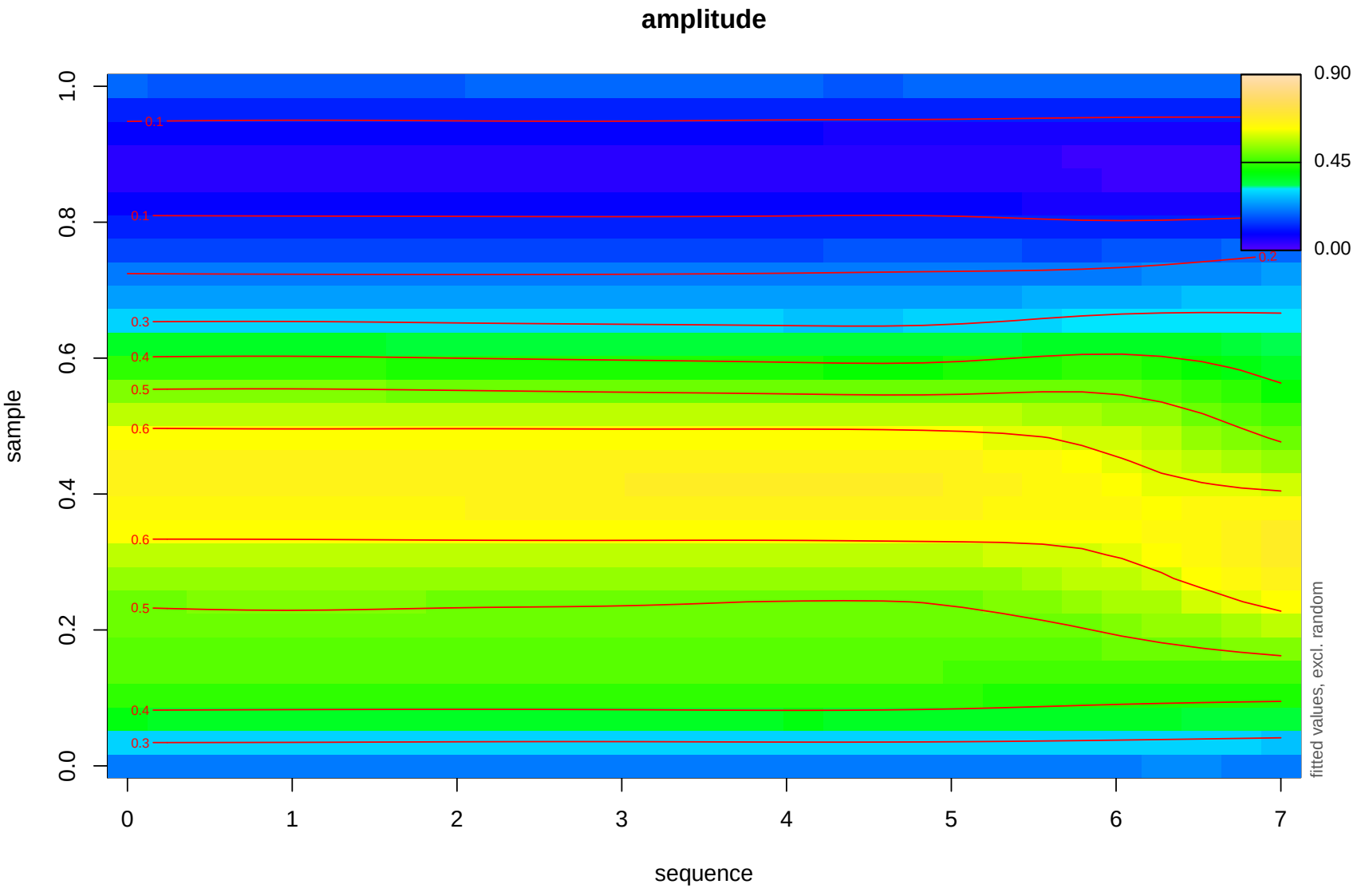
Results: Wavegram GAM

- generalised additive mixed models (Wood, 2006; Sóskuthy, 2017; van Rij et al., 2017)
    - non-linear multidimensional data
- statistical testing of wavegram data
    - heat-map plots: time, period, amplitude

Results: Wavegram GAM (modal voice)



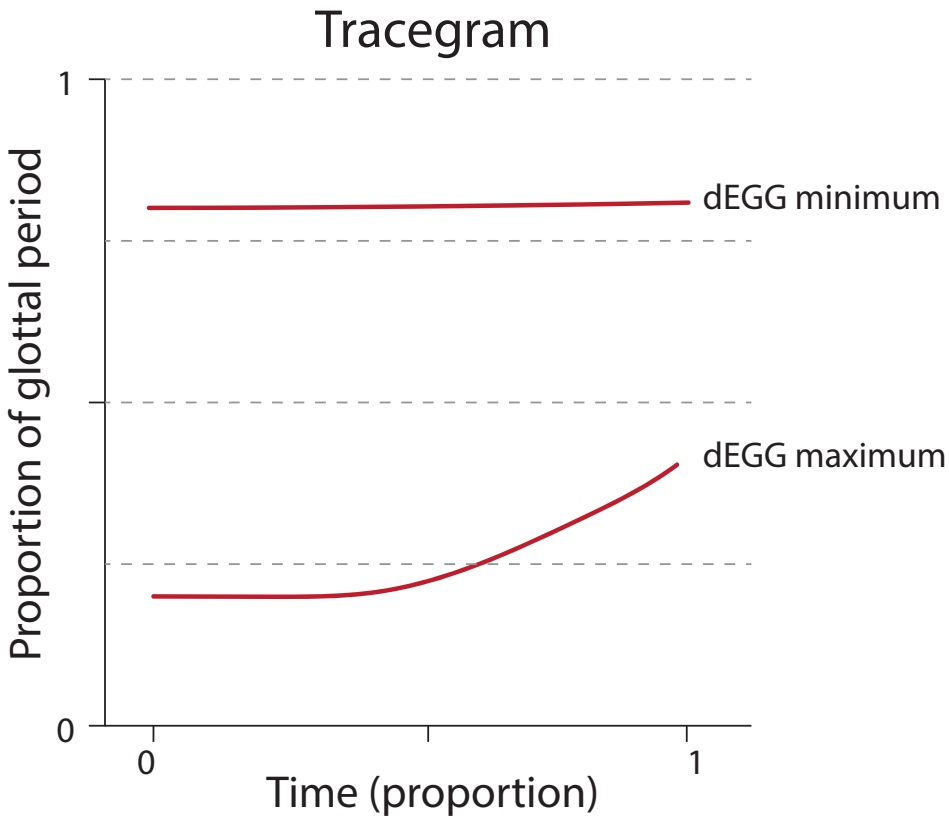
Results: Wavegram GAM (breathy voice)



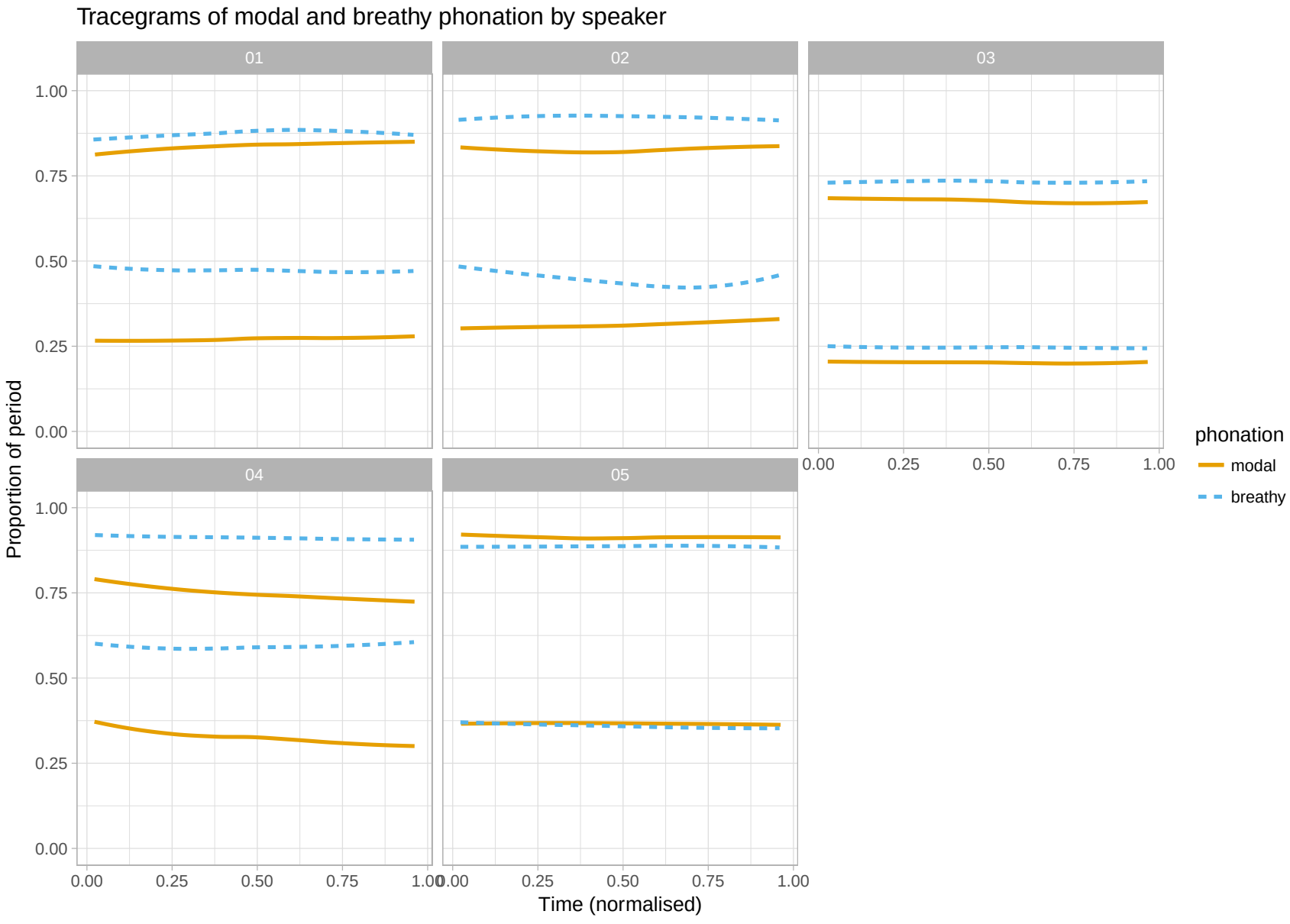
Results: Wavegram GAM

```
## phonation_gam_null: amplitude ~ s(sequence, k = 8) + s(sample) + ti(sequence, sample,
##     k = 8) + s(sequence, speaker_phon, bs = "fs", m = 1, k = 8)
##
## phonation_gam: amplitude ~ phonation + s(sequence, k = 8) + s(sample) + s(sequence,
##     by = phonation, k = 8) + s(sample, by = phonation) + ti(sequence,
##     sample, k = 8) + ti(sequence, sample, by = phonation, k = 8) +
##     s(sequence, speaker_phon, bs = "fs", m = 1, k = 8)
##
## Chi-square test of ML scores
## -----
##                Model     Score Edf Difference    Df  p.value Sig.
## 1 phonation_gam_null -53190.37  10
## 2      phonation_gam -66983.42  18  13793.050 8.000  < 2e-16  ***
##
## AIC difference: 27741.14, model phonation_gam has lower AIC.
```

Results: Tracegram



Results: Tracegram



Discussion

- CQ performed badly for speaker 03
- Tracegrams
    - non-resource-intensive method for visualising fold activity
- Wavegram GAMs
    - assessing fold activity data statistically

The end

Thanks!

References

Awan, Shaheen N., Andrew R. Krauss & Christian T. Herbst. 2015. An examination of the relationship between electroglottographic contact quotient, electroglottographic decontacting phase profile, and acoustical spectral moments. *Journal of Voice* 29(5). 519–529. doi:10.1016/j.jvoice.2014.10.016.

Baken, Ronald J. 1992. Electroglottography. *Journal of Voice* 6(2). 98–110.

Fabre, P. 1957. Un procede electrique percutane d'inscrition de l'accolement glottique au cours de la phonation: glottographie de haute frequence. Premiers resultats. *Bulletin de l'Académie nationale de médecine* 141. 66.

Hampala, Vít, Maxime Garcia, Jan G. Švec, Ronald C. Scherer & Christian T. Herbst. 2016. Relationship between the electroglottographic signal and vocal fold contact area. *Journal of Voice* 30(2). 161–171. doi:10.1016/j.jvoice.2015.03.018.

Herbst, Christian T., W. Tecumseh S. Fitch & Jan G. Švec. 2010. Electroglottographic wavegrams: A technique for visualizing vocal fold dynamics noninvasively. *The Journal of the Acoustical Society of America* 128(5). 3070–3078. doi:10.1121/1.3493423.

Herbst, Christian T., Jörg Lohscheller, Jan G. Švec, Nathalie Henrich, Gerald Weissengruber & W. Tecumseh S. Fitch. 2014. Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings. *Journal of Experimental Biology* 217(6). 955–963. doi:10.1242/jeb.093203.

Herbst, Christian T., Harm K. Schutte, Daniel L. Bowling & Jan G. Svec. 2017. Comparing chalk with cheese—the EGG contact quotient is only a limited surrogate of the closed quotient. *Journal of Voice* 31(4). 401–409.

Rothenberg, Martin & James J. Mahshie. 1988. Monitoring vocal fold abduction through vocal fold contact area. *Journal of Speech, Language, and Hearing Research* 31(3). 338–351.

Scherer, Ronald C. & Ingo R. Titze. 1987. The abduction quotient related to vocal quality. *Journal of Voice* 1(3). 246–251.

Sóskuthy, Márton. 2017. Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. arXiv preprint arXiv:1703.05339.

van Rij, Jacolien, Martijn Wieling, R. Harald Baayen & Hedderik van Rijn. 2017. itsadug: Interpreting time series and autocorrelated data using GAMMs. R package version 2.3.

Wood, Simon. 2006. *Generalized additive models: An introduction with R*. CRC Press.