Analysing fixations in a visual-world eye tracking paradigm – Why GAMMs are a good choice!

CLaS Eye-tracking Workshop 2: Data analysis and statistical approaches
Macquarie University, Sydney, Thursday 8th September 2022

Katharina Zahner-Ritter
Phonetics, University of Trier

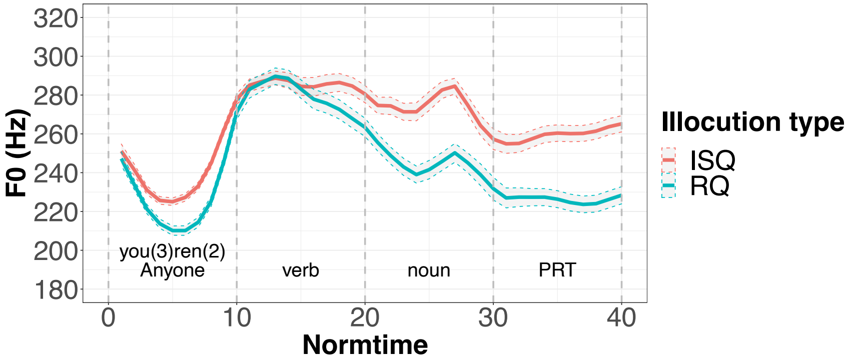
UNIVERSITÄT TRIER

Phonetics Lab TRIER UNIVERSITY

My research foci & and why GAMMs* are a helpful companion

* Generalized Additive Mixed Models

Production: F0 contours (over time)

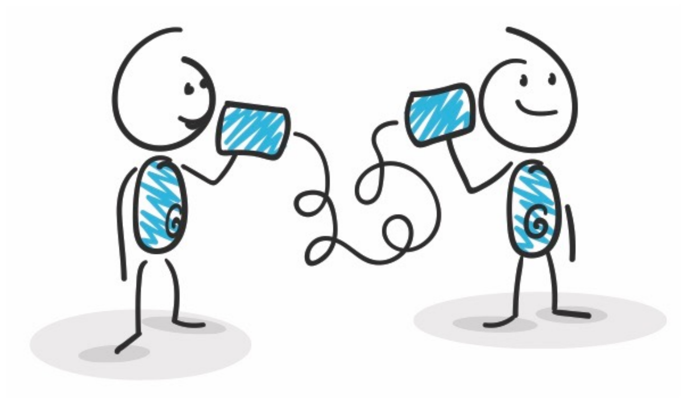


有人(yǒurén) 吃(chī) 柠檬(níngméng) 么(me)?

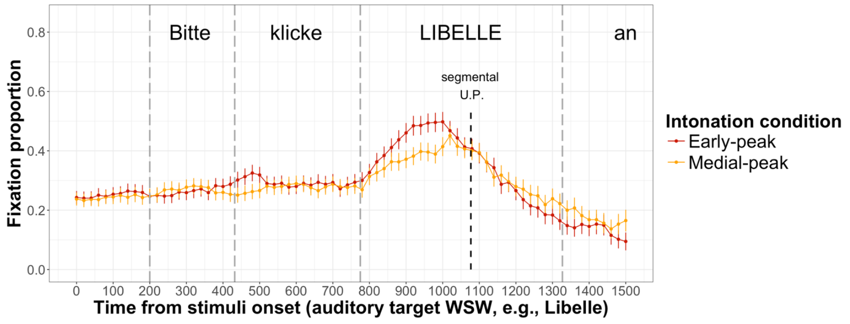
Subject	**verb**	**noun**	**PRT**
Anybody	eat	lemon	PRT

"Does anyone eat lemon?"

(Zahner-Ritter et al. accepted, JPhon; Zahner-Ritter et al. 2022a, Frontiers Comm, Zahner-Ritter et al. 2022b, Frontiers Psy)



Processing: Fixations (over time)



(Zahner et al. 2019, JPhon;
Zahner-Ritter et al. 2022, SLE)

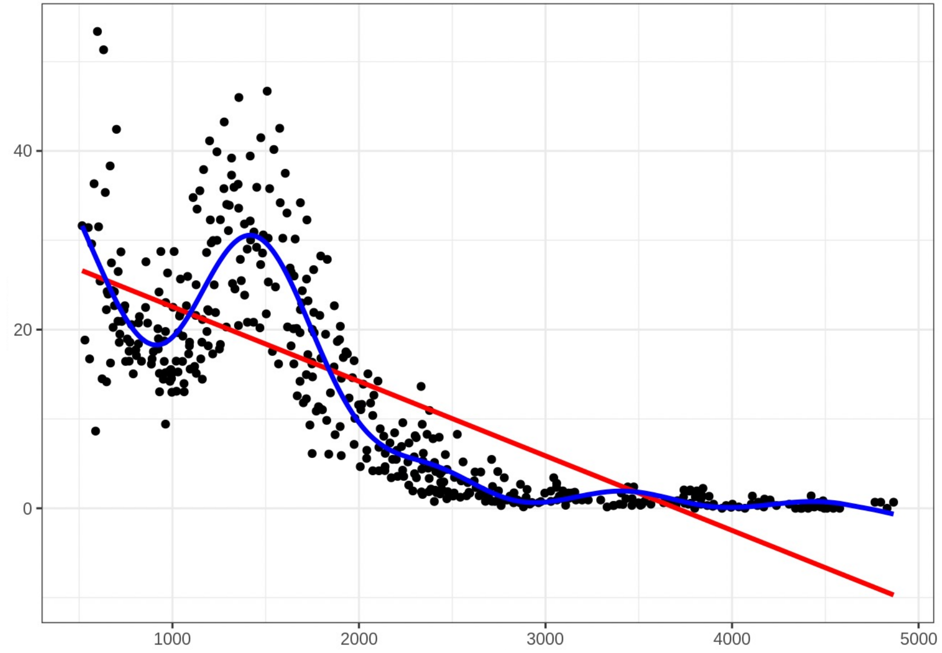
Generalized Additive Mixed Models (GAMMs) in a nutshell

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

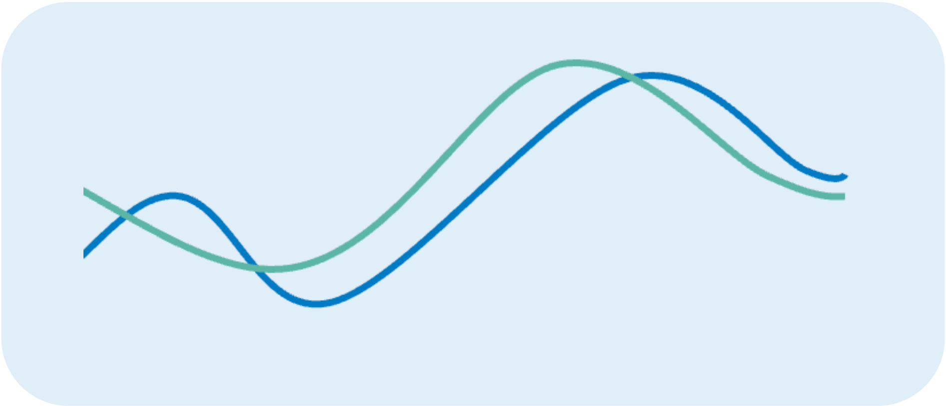
In a nutshell: Why GAMMs

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

- Suited to analyze nonlinear time-series data (f0, fixations, EEG etc.)



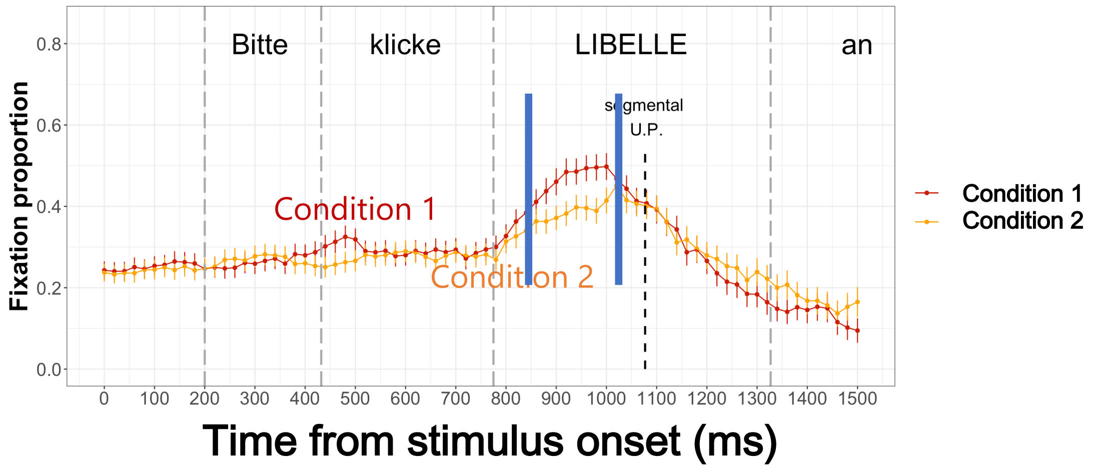
http://r.qcbs.ca/workshop08/book-en/introduction-to-gams.html

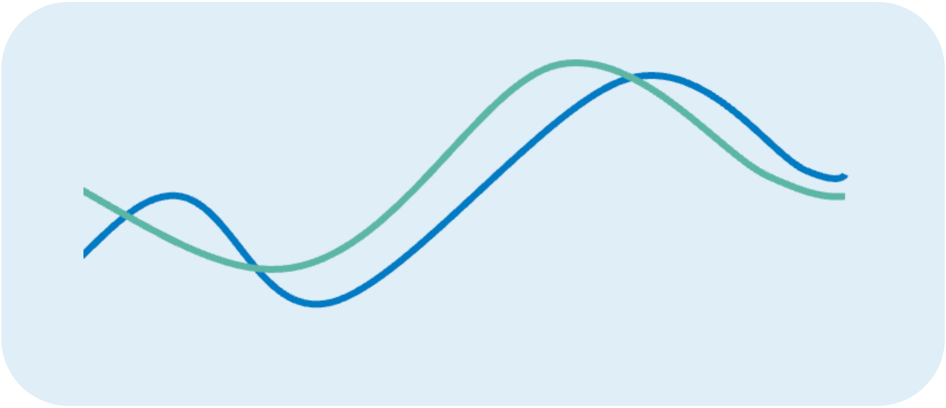


In a nutshell: Why GAMMs

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

- Suited to analyze nonlinear time-series data (f0, fixations, EEG etc.)
- Allow for a "holistic" analysis of the time-series data
  → No averaging over (arbitrary) time windows necessary (e.g., fixations)
  → More informative, e.g., as to the shape of the contour (e.g., f0)
- Allow us to study when in time an effect arises

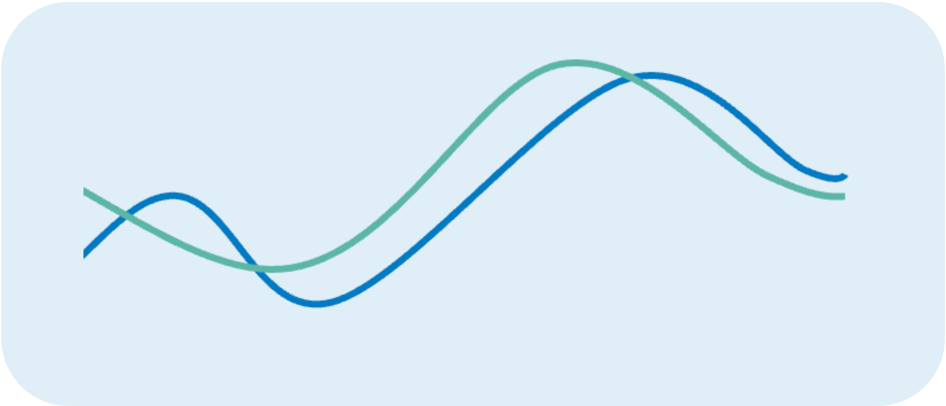




In a nutshell: Why GAMMs

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

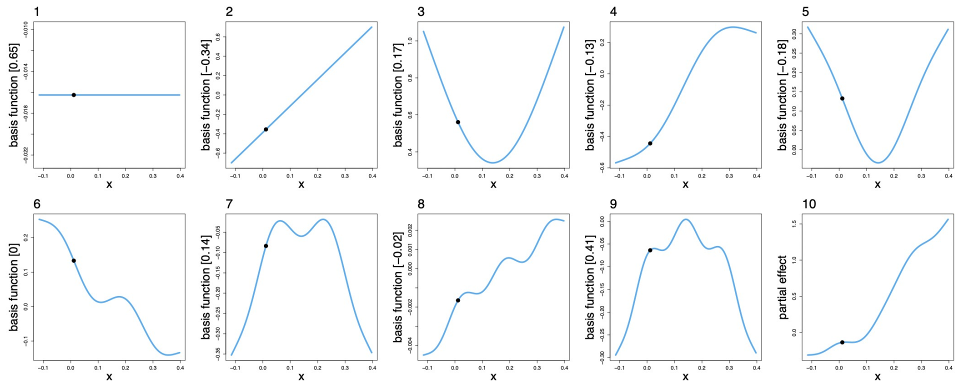
- Suited to analyze nonlinear time-series data (f0, fixations, EEG etc.)
- Allow for a "holistic" analysis of the time-series data
  → No averaging over (arbitrary) time windows necessary (e.g., fixations)
  → More informative, e.g., as to the shape of the contour (e.g., f0)
- Allow us to study when in time an effect arises
- Models can account for autocorrelation in time-series data (Gaussian models); otherwise overestimation of effect
- Inclusion of complex
  - (non-linear) interactions between variables
  - (non-linear) random effects
  - control variables



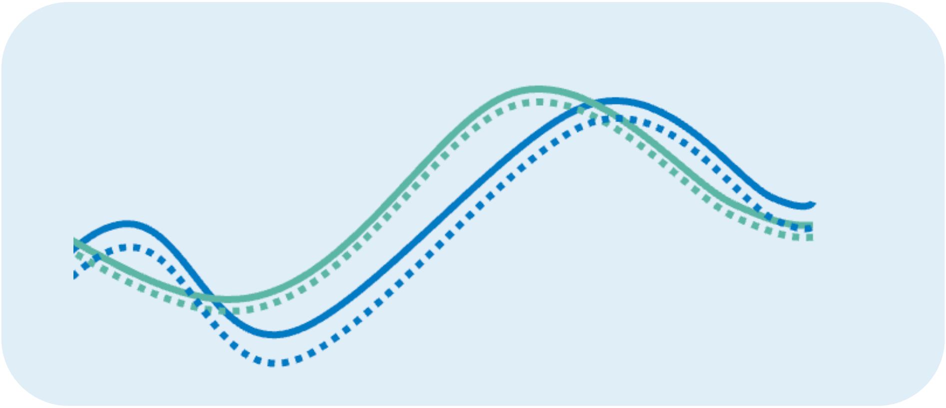
In a nutshell: How GAMMs work

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

- Model looks for the nonlinear function that accounts for the raw data best; (solid line represents raw data; dotted line predictions)
- In simple words, it adds up a pre-specified number of basis functions of different shapes to result in final curve:



(Baayen & Linke, in press, p. 7)

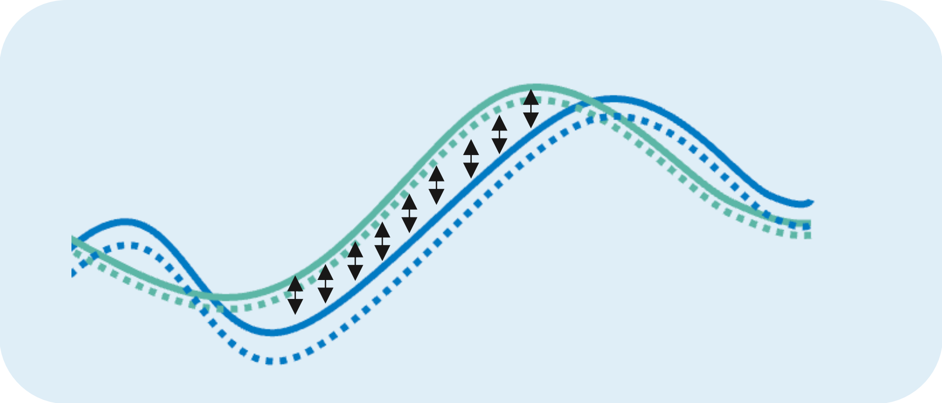


→ Basis functions weighted on how well they fit data
(i.e., the magnitude of the contribution of that basis function)

In a nutshell: GAMM results

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

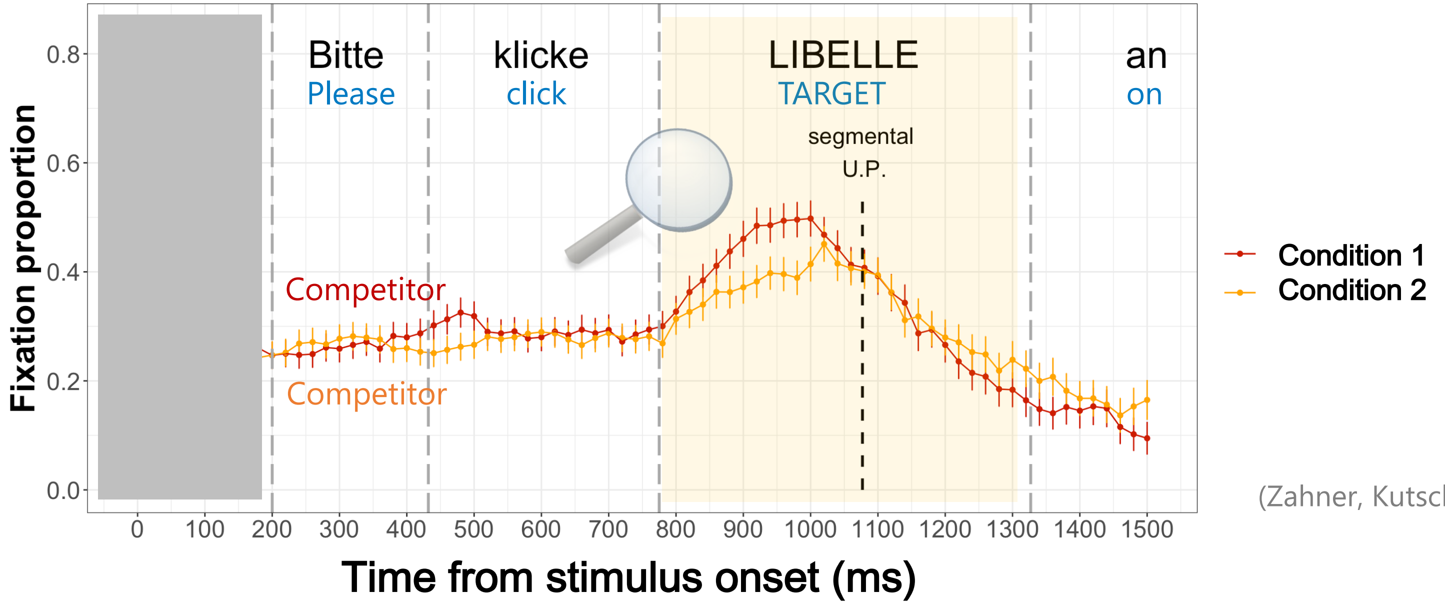
- GAMM results are most intuitively interpreted via visualization, typically so called difference curves
- Difference curves show the predicted difference between two conditions over time (e.g., green minus blue dotted line)
- Difference curves reveal when in time two conditions significantly differ



In a nutshell: GAMM difference curves

(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)

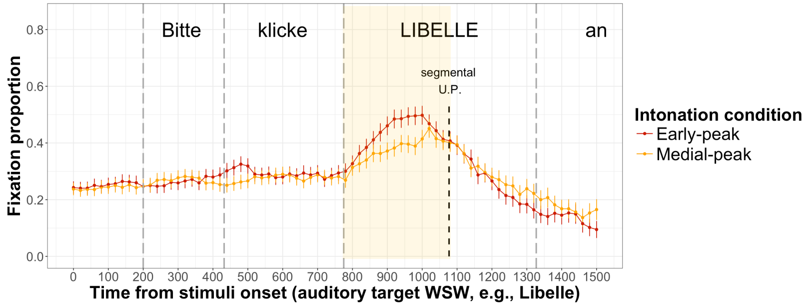
Example for Visual-World Eye-Tracking:
Are there more **competitor fixations** in **condition 1** compared to **condition 2**?



(Zahner, Kutscheid, Braun, 2019, JPhon)

In a nutshell: GAMM difference curves

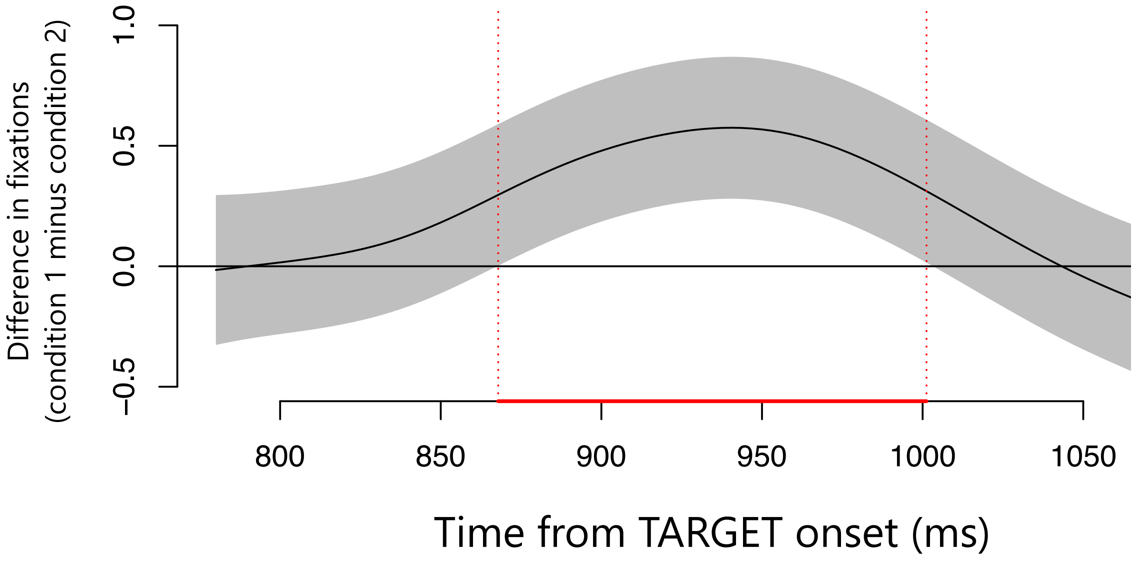
(Baayen et al., 2018; Wieling, 2018; Wood, 2006; Wood, 2017)



Example for Visual-World Eye-Tracking:

Are there more **competitor fixations** in **condition 1** compared to **condition 2**?

(Zahner, Kutscheid, Braun, 2019, JPhon)



More competitor fixations
in condition 1

0 = No difference

More competitor fixations
in condition 2

Grey band indicates the 95% CI of the **mean difference**

Red vertical lines indicate significant difference: 868-1001 ms

UNIVERSITÄT TRIER

Helpful resources & tutorials (non-exhaustive list)

Many thanks to Harald Baayen, Jacolien van Rij, Vincence Porretta & Cecsko Voeten for great GAMM workshops!

- Baayen, R. H., and Linke, M. (2020). An introduction to the generalized additive model. In Gries, S. Th. and M. Paquot (Eds.) A practical handbook of corpus linguistics. Springer, Berlin. http://www.sfs.uni-tuebingen.de/~hbaayen/publications/BaayenLinke2020.pdf
- Porretta, V., Kyröläinen, A.-J., van Rij, J., & Järvikivi, J. (2018). Visual world paradigm data: From preprocessing to nonlinear time-course analysis. In I. Czarnowski, R. Howlett & L. Jain (Eds.), Intelligent Decision Technologies 2017, number 73 (pp. 268-277).
- Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. Journal of Phonetics, 58, 1-21.
- Sóskuthy, M. (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. Journal of Phonetics, 84.
- van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., & Wood, Simon N. (2019). Analyzing the time course of pupillometric data. Trends in Hearing, 23, 1-22.
- Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. Journal of Phonetics, 70, 86-116. http://www.let.rug.nl/wieling/Tutorial).
- http://r.qcbs.ca/workshop08/book-en/

Analyzing fixations in a VW Paradigm using GAMMs – An example study

Zahner, K., Kutscheid, S., & Braun, B. (2019). Alignment of f0 peak in different pitch accent types affects perception of metrical stress. *Journal of Phonetics*, 74, 75-95.

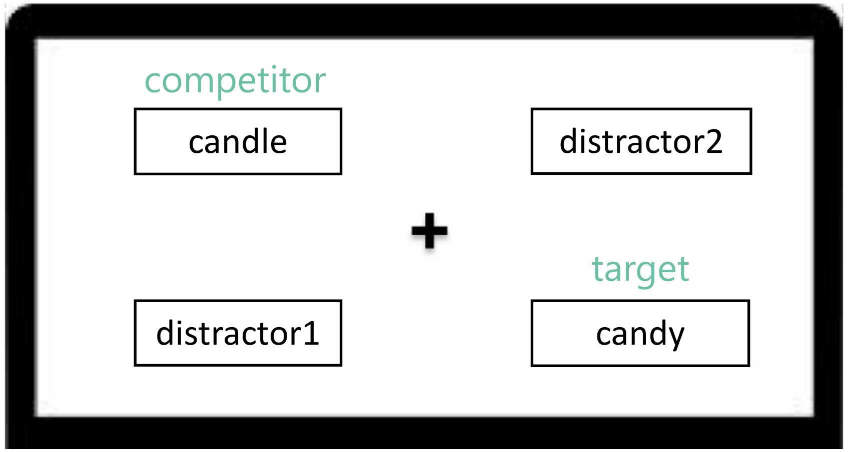
https://data.mendeley.com/datasets/2gkpwpg44j/3

Other studies that have used GAMMs for analysing VW data
(e.g., Nixon, van Rij, Mok, Baayen, & Chen, 2016; Porretta, Tucker, & Järvikivi, 2016;
van Rij, Hollebrandse, & Hendriks, 2016).

The Visual-World Eye-Tracking Paradigm (VW Paradigm)

(Tanenhaus et al., 1995; McQueen & Viebahn, 2007; Reinisch et al., 2010; Huettig et al., 2011 for a review)

- Tracking of fixations to targets (words, pictures) on screen while presenting acoustic stimuli (e.g., Please click on <target>)
- Fixations as a "window to the brain":
  Listeners direct fixations to the visual representations they are currently processing as the stimulus unfolds (online processing)
- Uncovering the dynamics of spoken language processing
- Crucial: Time-locking between fixations and acoustic stimulus



Audio: "Please click on candy"

The Visual-World Eye-Tracking Paradigm (VW Paradigm)

(Tanenhaus et al., 1995; McQueen & Viebahn, 2007; Reinisch et al., 2010; Huettig et al., 2011 for a review)

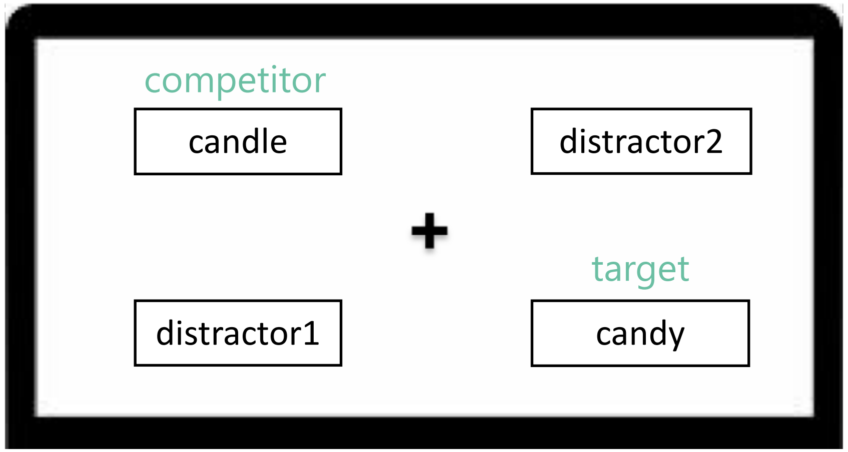
- Evidence from VW Paradigm converges on the fact that listeners use both

  segmental (candle, candy) and

  suprasegmental cues (octopus, October)

  as soon as they become available.

  (Allopenna et al., 1998; McQueen & Viebahn, 2007; Connell et al., 2018; Tanenhaus et al., 1995; Jesse et al., 2017; Reinisch et al., 2010; Sulpizio & McQueen, 2012)



Audio: “Please click on candy”

Example study

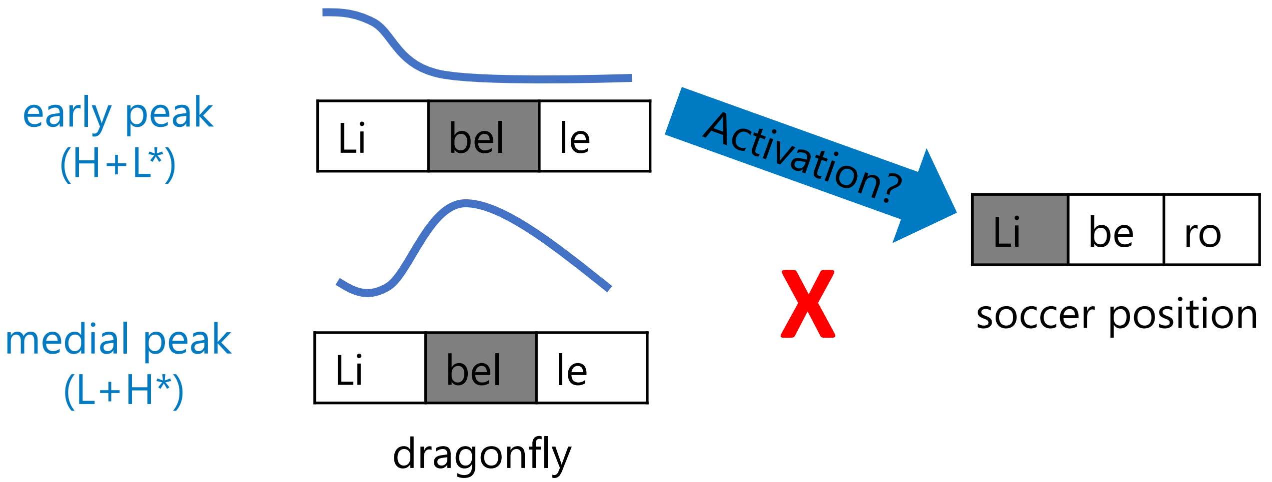
Zahner, K., Kutscheid, S., & Braun, B. (2019). Alignment of f0 peak in different pitch accent types affects perception of metrical stress. *Journal of Phonetics*, 74, 75-95.

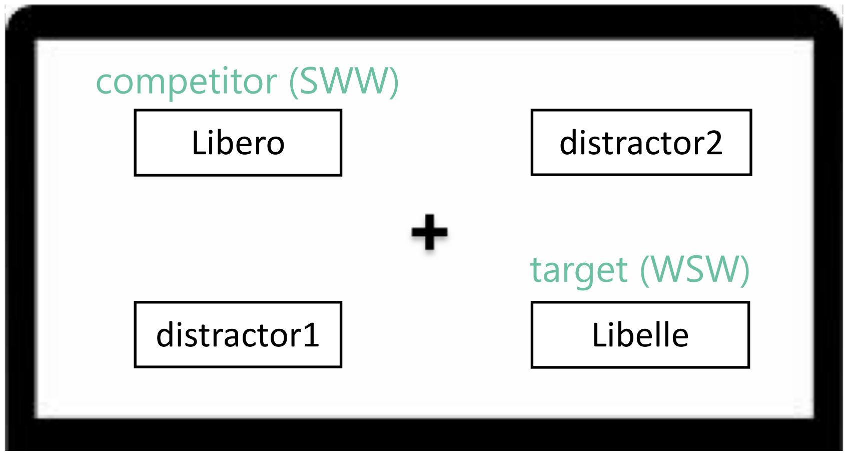
https://data.mendeley.com/datasets/2gkpwpg44j/3

Rationale of the study

Research question

- Do f0 peaks on unstressed initial syllables in a WSW word lead to temporary lexical activation of cohort competitors with initial stress (SWW) in online processing?





Audio:
“Bitte klicke Libelle an.”

(intonation condition was manipulated:
early- vs. medial peak accent)

Materials

Target (WSW)	Stress Competitor (SWW)
Libelle [liˈbɛlə]	Libero [ˈlibəro]
Kaverne [kaˈvɛɐnə]	Kaviar [ˈkaviaɐ]
Albaner [alˈbanɐ]	Albatros [ˈalbatrɔs]

…

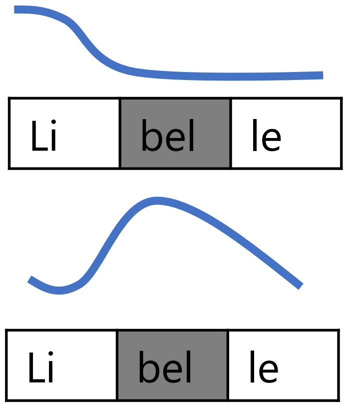
"Bitte klicke Libelle an".

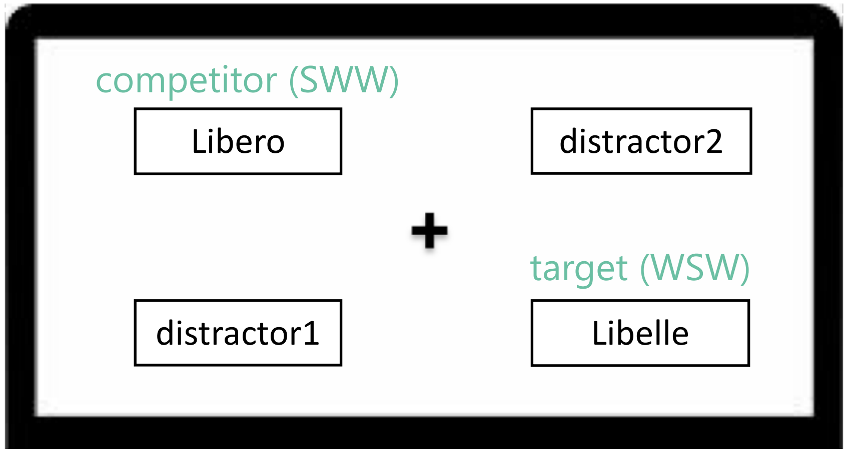
<WSW target>

'Please click on <Target>'

early peak
(H+L*)

medial peak
(L+H*)





Audio:
"Bitte klicke Libelle an."

(intonation condition was manipulated;
Latin Square)

(Trial exported with Data Viewer; data from Zahner et al. 2019)

Preparing data for statistical analyses

VWPre: Tools for Preprocessing Visual World Data

Gaze data from the Visual World Paradigm requires significant preprocessing prior to plotting and analyzing the data. This package provides functions for preparing visual world eye-tracking data for statistical analysis and plotting. It can prepare data for linear analyses (e.g., ANOVA, Gaussian-family LMER, Gaussian-family GAMM) as well as logistic analyses (e.g., binomial-family LMER and binomial-family GAMM). Additionally, it contains various plotting functions for creating grand average and conditional average plots. See the vignette for samples of the functionality. Currently, the functions in this package are designed for handling data collected with SR Research Eyelink eye trackers using Sample Reports created in SR Research Data Viewer. While we would like to add functionality for data collected with other systems in the future, the current package is considered to be feature-complete; further updates will mainly entail maintenance and the addition of minor functionality.

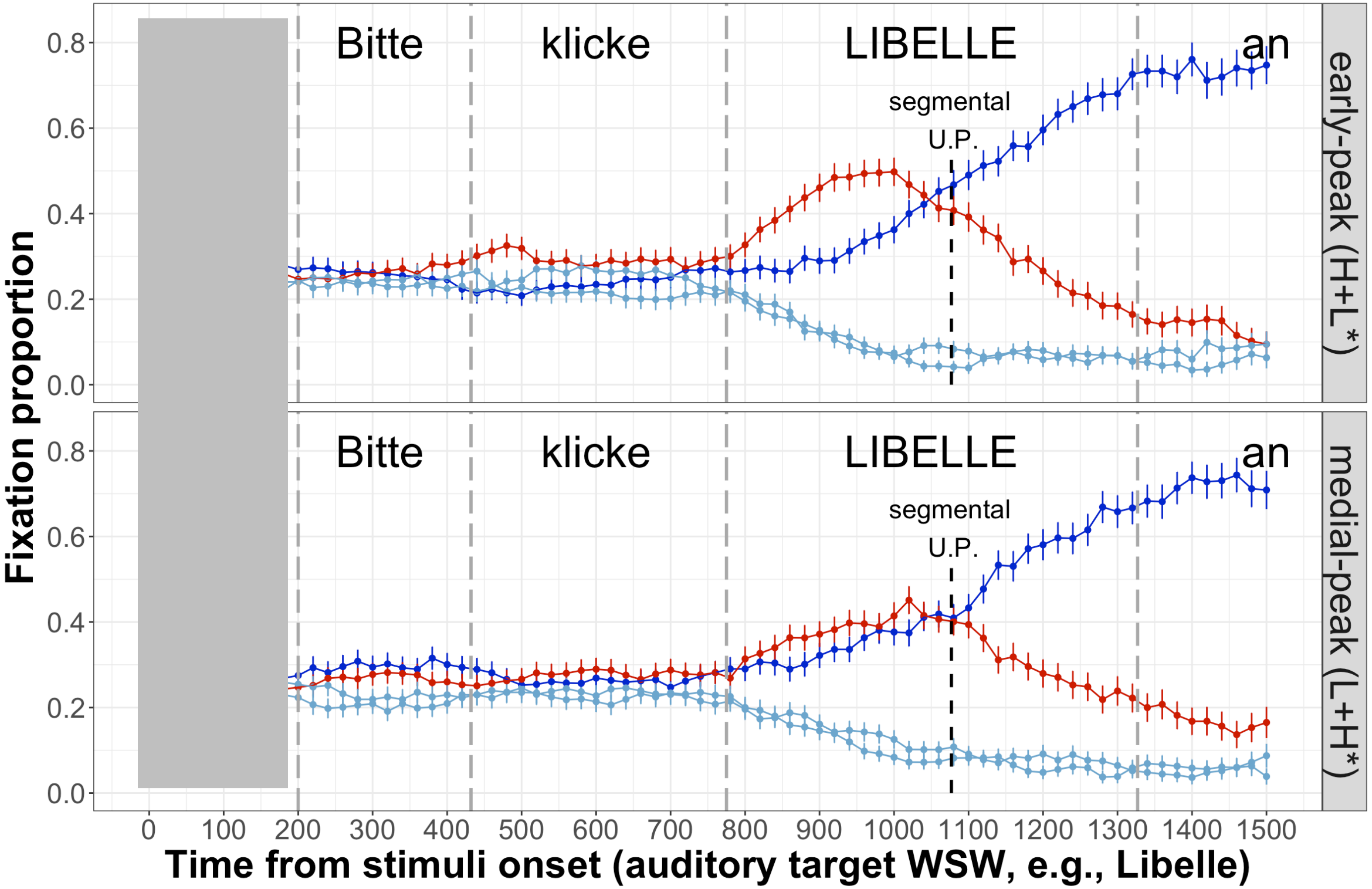
Vignettes –
highly recommended!

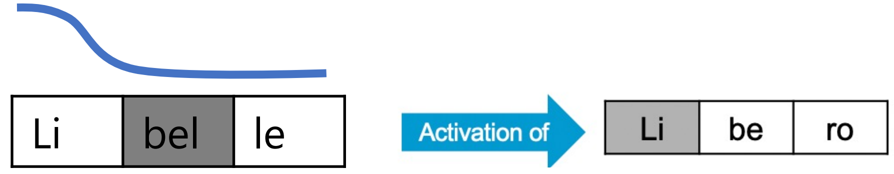
[Basic VWP Preprocessing]
[Relabeling or Defining Interest Areas]
[Aligning Data to a Specific Sample Message]
[Plotting VWP Data Processed with VWPre]

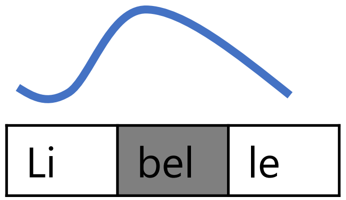
Porretta, V., Kyröläinen, A., van Rij, J., & Järvikivi, J. (2017). VWPre: Tools for preprocessing visual world data, R package.

https://cran.r-project.org/web/packages/VWPre/index.html

Visualizing data: Evolution of fixations to four words on screen (areas of interest)



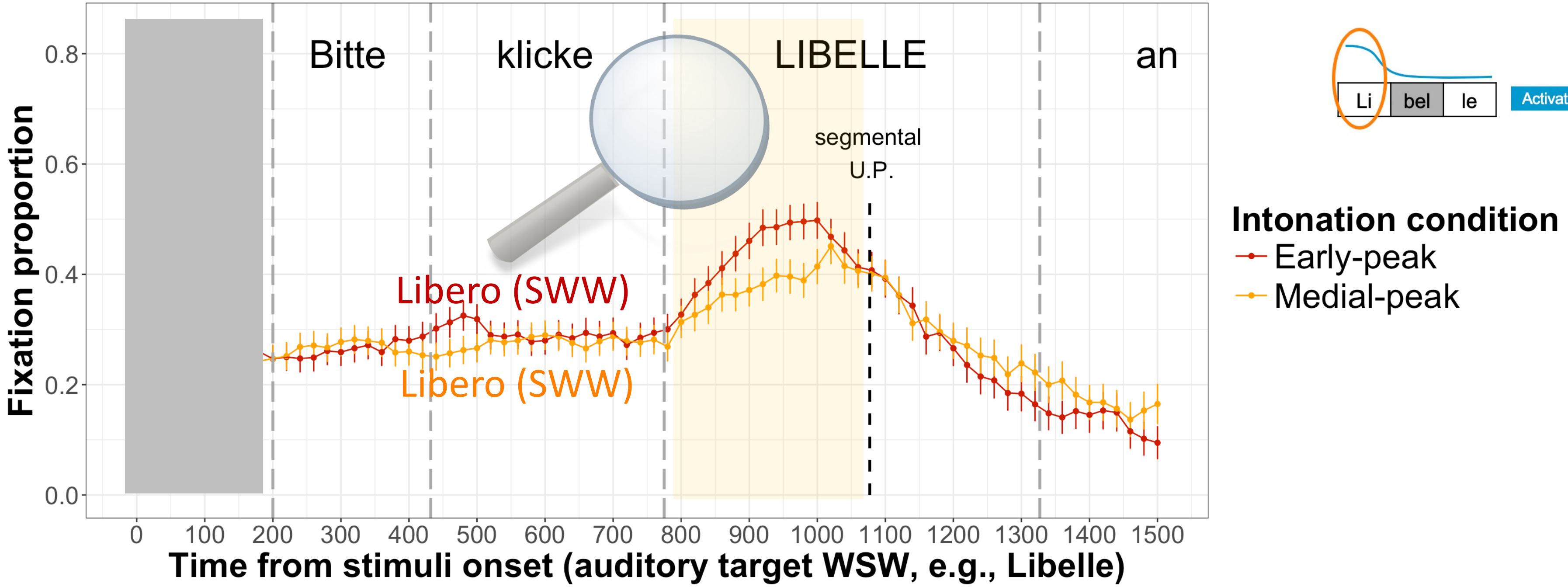




(Zahner et al. 2019)

 https://cran.r-project.org/web/packages/VWPre/vignettes/VWPre_Plotting.html

Comparison of SWW competitor fixations over time



Analysis: GAMMs

Some (basic) comments on modelling

- The analysis was done in R (R Development Core Team, 2015, closely following Wieling 2018 and Porretta et al. 2016)
- Code available on Mendeley: https://data.mendeley.com/drafts/2gkpwpg44j

Packages

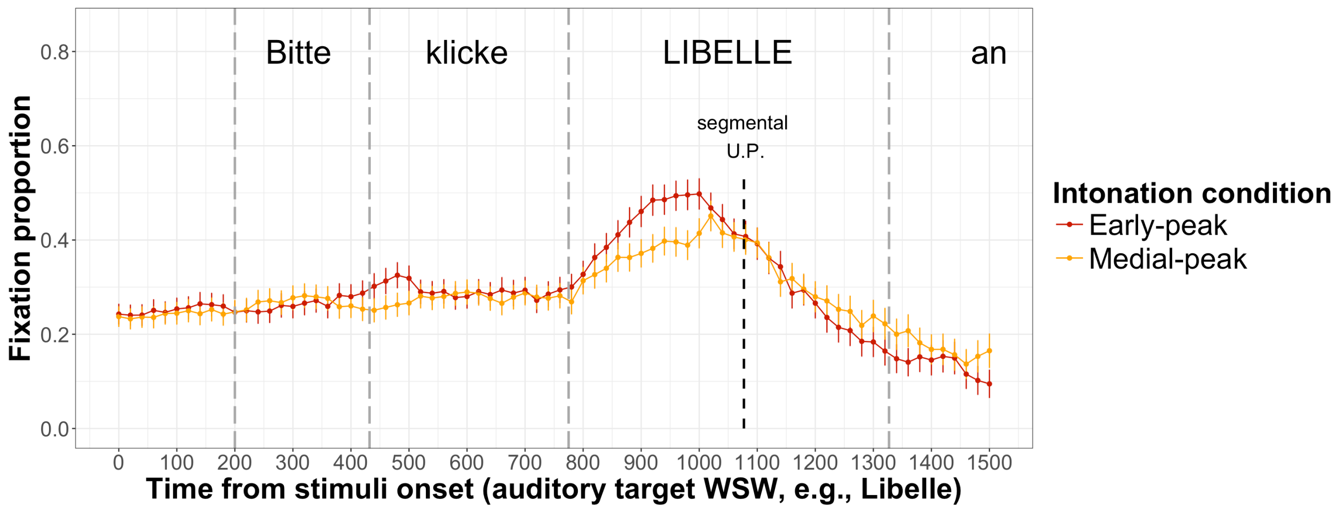
```
library(VWPre)     # preprocessing of VW data (see Porretta et al. 2017)
library(ggplot2)   # for visualizing raw data (fixation plots)
library(mgcv)      # for GAMM modelling (see Wood, 2011, 2017)
library(itsadug)   # for GAMM visualization (see van Rij, Wieling, Baayen, & van Rijn, 2017)
```

Basic Model: Competitor fixations

```
comp.fix <- bam(IA_2_ELogit ~               # response variable: elogs (Barr, 2008)
                cond +                      # parametric effect of intonation condition
                s(TIMESTAMP, by = cond)     # cond-smooth; nonlinear functional relations with the DV over time
                s(event, bs = 're'),        # random intercept for event, comb. item and subject as unique identifier
                data = data,                # dataset
                method = "ML")              # Maximum Likelihood
```

Response variable: Competitor fixations
empirical logits (elogs): logit transformed ratio of the fixations to the competitor divided by the fixations directed to the three other objects (Barr, 2008)

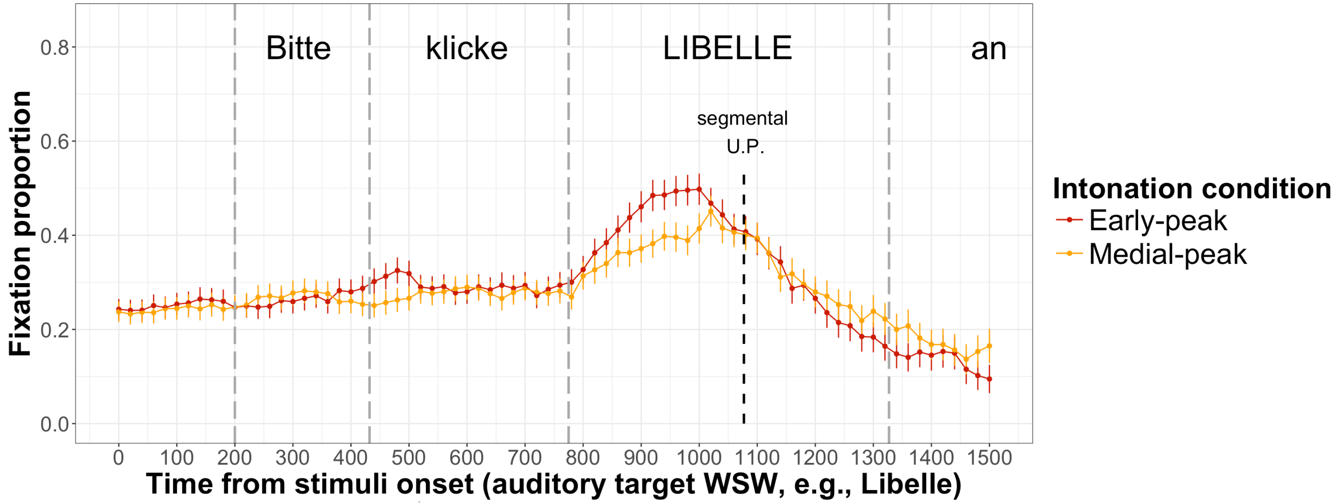
Proportions bound between 0-1;
Logits provide a transformation resulting in an unbounded measure



Basic Model: Competitor fixations

```
comp.fix <- bam(IA_2_ELogit ~             # response variable: elogs (Barr, 2008)
                cond +                    # parametric effect of intonation condition
                s(TIMESTAMP, by = cond) # cond-smooth; nonlinear functional relations with the DV over time
                s(event, bs = 're'),      # random intercept for event, comb. item and subject as unique identifier
                data = data,              # dataset
                method = "ML")            # Maximum Likelihood
```

```
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.4035     0.1012  -3.988 6.74e-05 ***
## condm        -0.2285     0.1406  -1.626    0.104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                         edf  Ref.df      F p-value
## s(TIMESTAMP):conde    5.416   6.718 30.075  <2e-16 ***
## s(TIMESTAMP):condm    1.688   2.133 44.491  <2e-16 ***
## s(event)            643.277 730.000  8.655  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.648   Deviance explained = 67.7%
## -ML =  11495  Scale est. = 1.4175     n = 7917
```



Basic Model: Competitor fixations

```
comp.fix <- bam(IA_2_ELogit ~                 # response variable: elogs (Barr, 2008)
                cond +                        # parametric effect of intonation condition
                s(TIMESTAMP, by = cond) # cond-smooth; nonlinear functional relations with the DV over time
                s(event, bs = 're'),          # random intercept for event, comb. item and subject as unique identifier
                data = data,                  # dataset
                method = "ML")                # Maximum Likelihood
```

Accounting for autocorrelation in data

Additional To-Dos

```
acf_resid()
```

Checking model assumptions

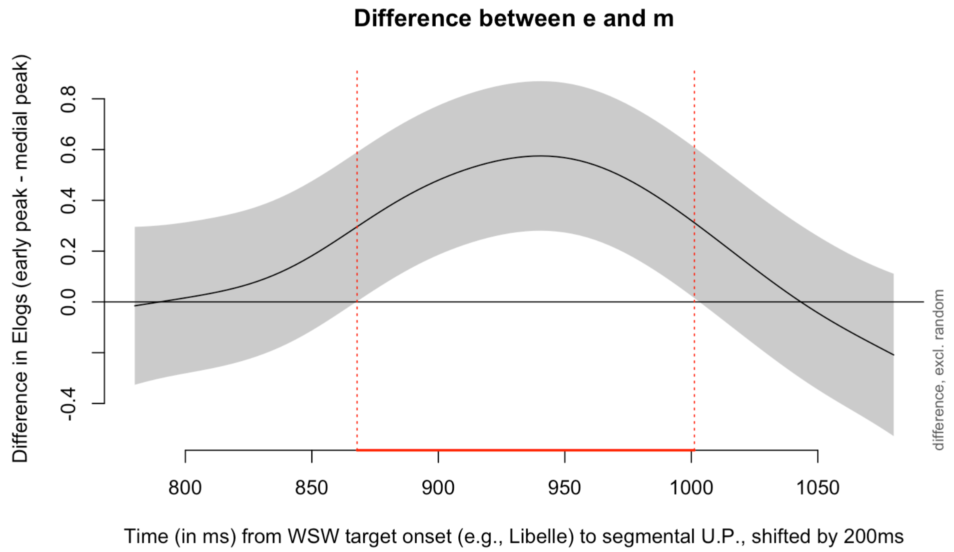
```
gam.check(model)
```

Model comparisons

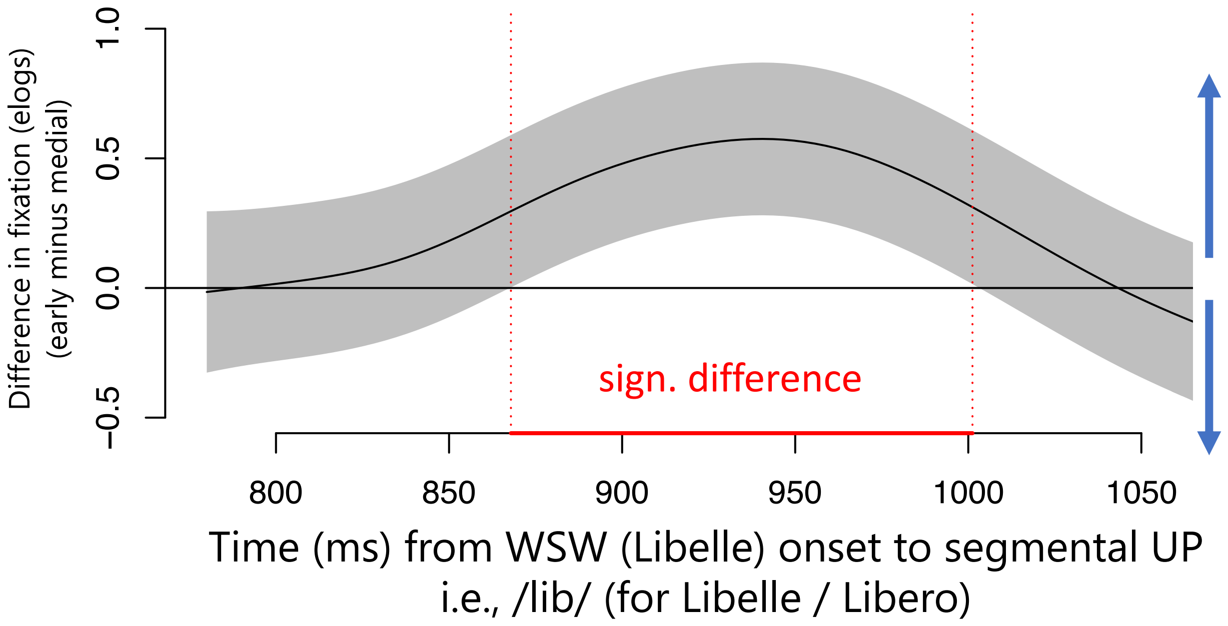
```
compareML(model1, model2)
```

Difference curve: competitor fixations

```
plot_diff(comp.fix,
          view = "TIMESTAMP",
          comp = list(cond = (c("e","m"))),
          ylab = "Difference in Elogs (early peak - medial peak)",
          xlab = "Time (in ms) from WSW target onset (e.g., Libelle) to segmental U.P., shifted by 200ms",
          rm.ranef = T)
```



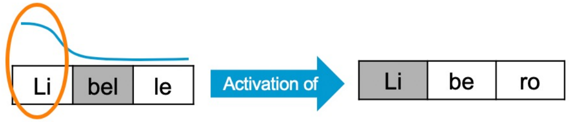
Interpreting GAMM difference curves – Competitor fixations



More competitor (Libero) fixations in early-peak condition

0 = No difference

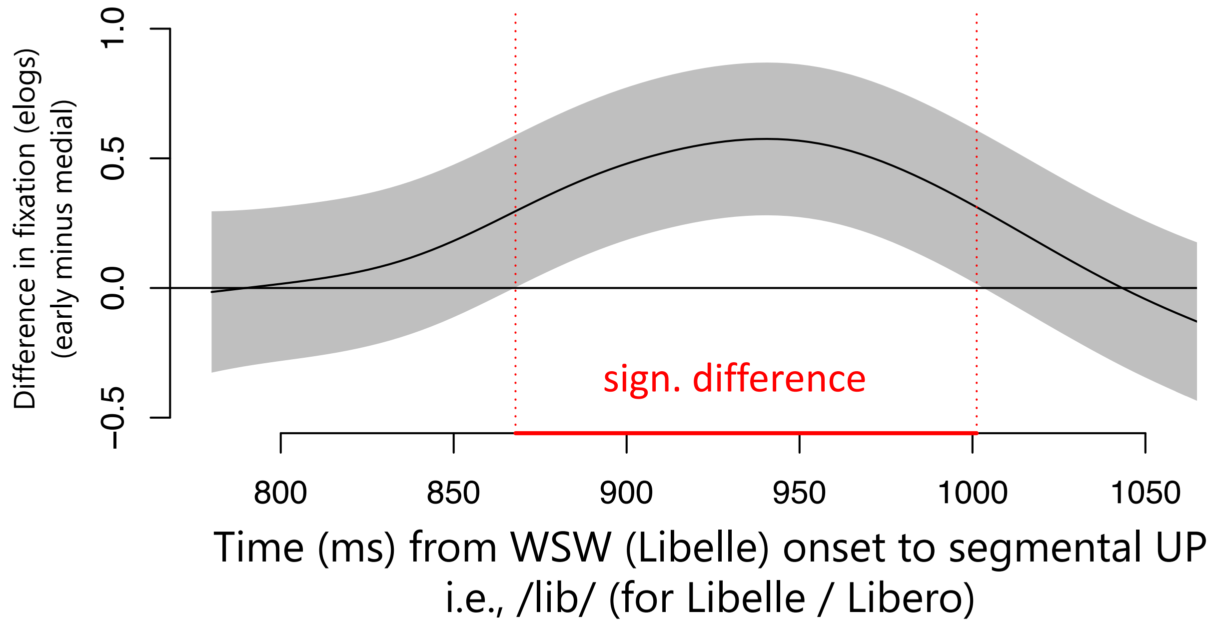
Fewer competitor (Libero) fixations in early-peak condition



Effect of intonation condition

Recall: Grey band indicates the 95% CI of the mean difference
Red lines indicate significant difference: 868-1001 ms

Interpreting GAMM difference curves – Competitor fixations



Recall: Grey band indicates the 95% CI of the mean difference
Red lines indicate significant difference: 868-1001 ms

Pitch accent type affects lexical activation in German adults

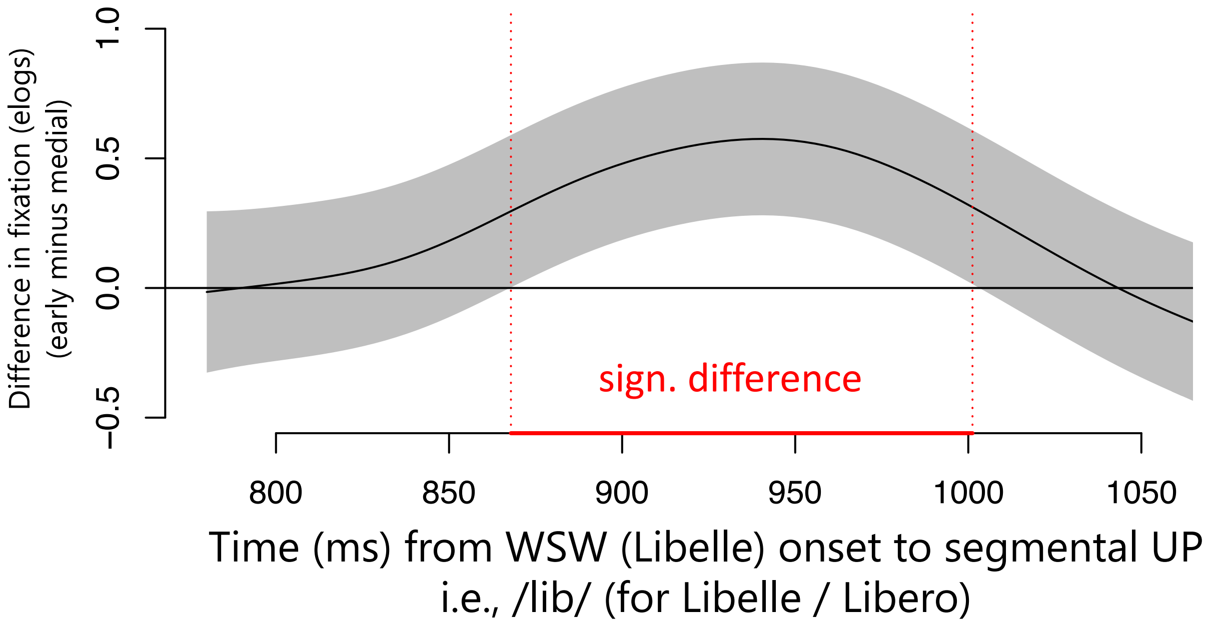
More fixations to competitor when target was realized with early-peak accent (H+L*), compared to when realized with a medial-peak accent (L+H*)

F0 peak in the form of a H-leading tone on unstressed syllable (H+L*) prompts percept of stressed syllable

Effect of intonation condition

Yes, but why?

Interpreting GAMM difference curves – Competitor fixations



Recall: Grey band indicates the 95% CI of the mean difference
Red lines indicate significant difference: 868-1001 ms

Effect of intonation condition

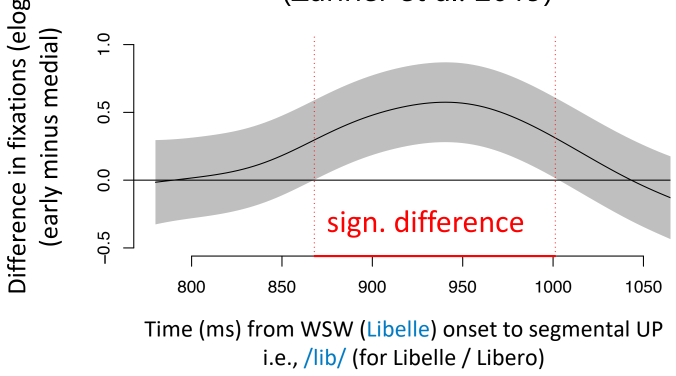
Yes, but why?

Salience or frequency?

Manipulation of occurrence frequency of H* / L*
in the immediate input
(prefixed exposure phases)

Results: GAMM difference curves – Competitor fixations

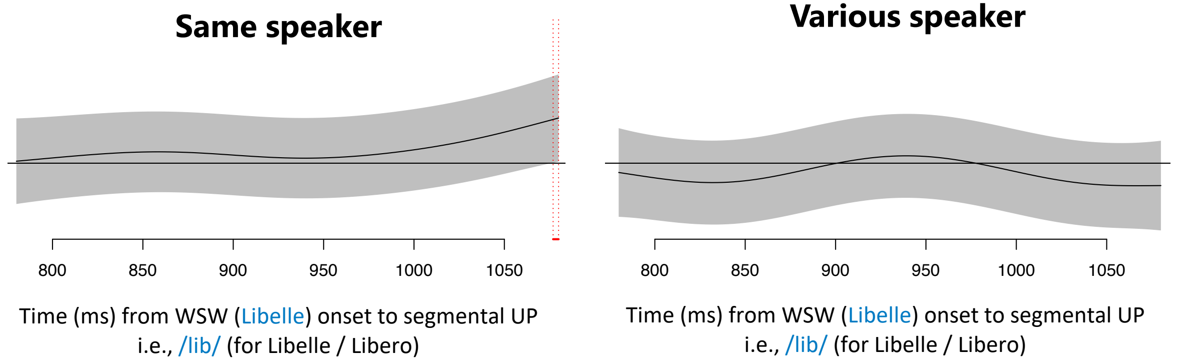
Without exposure phase
(Zahner et al. 2019)

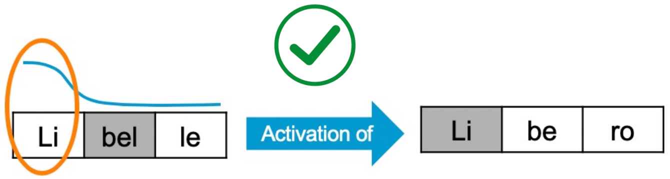


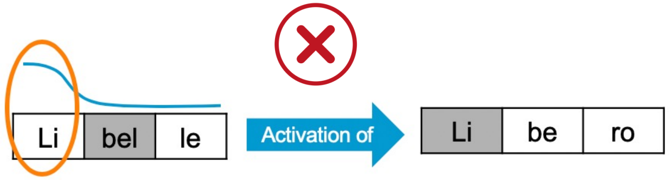
After exposure phase with **low-pitched accents**
(Zahner-Ritter et al. 2022, SLE)

Same speaker

Various speaker







Assessing interaction between experiment x condition (see Zahner et al. 2019; following analysis steps in Wieling, 2018,p. 106ff).

What we have not talked about today (but needs to be considered)

Accounting for autocorrelation in data

```
acf_resid()
```

Checking model assumptions

```
gam.check(model)
```

Model comparisons

```
compareML(model1, model2)
```

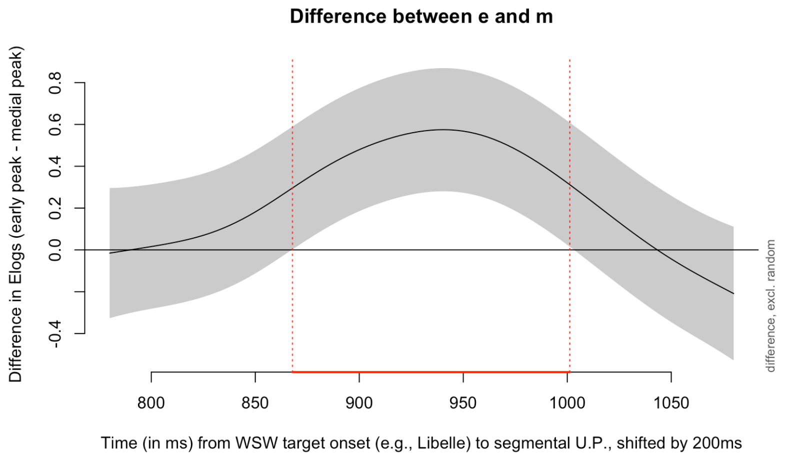
The whole topic of including:
- (non-linear) interactions between variables
- (non-linear) random effects
- control variables

Take-home message: Why GAMMs are a good choice to analyse fixations!

GAMMs represent a state-of-the-art statistical approach to analysing time-varying data with non-linear relationships and autocorrelation.

The visual representation of GAMMs indicates when in time an effect on a response variable becomes significant.

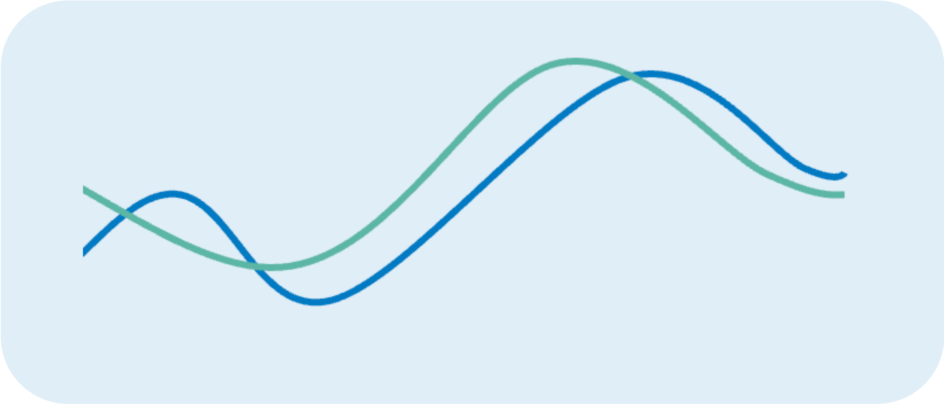
This is certainly an elegant alternative to traditional time-window analyses, which require fixations to be binned in predefined analyses windows (Barr, 2008), and perhaps also for other methods for analysing fixation data.



UNIVERSITÄT TRIER

Phonetics Lab
TRIER UNIVERSITY

Thank you very much for your attention!



Have fun in applying GAMMs to your data!

Special thanks to:

- Laurence Bruggeman for inviting me & and the organizing committee
- Bettina Braun for sharing statistical expertise
- Harald Baayen, Jacolien van Rij, Vincence Porretta & Cecsko Voeten for offering great GAMM workshops

Trier Center for Language and Communication
PATTERNS

Slides will be uploaded to:
https://katharinazahner.weebly.com

PHONFACTS_

zahnerritter@uni-trier.de

Twitter: @zahnerritter
Instagram: @phonfacts_

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419-439.

Arnhold, A., Porretta, V., Chen, A., Verstegen, S., Mok, I., Järvikivi, J., Iden, L. S. L. D. i. R. t. S. a., & Lapd, U. O. T. S. (2020). (Mis)understanding your native language: Regional accent impedes processing of information status. *Psychonomic Bulletin & Review*, 27, 801-808.

Baayen, R. H., & Linke, M. (2020). Generalized Additive Mixed Models. In M. Paquot & S. T. Gries (Eds.), A Practical Handbook of Corpus Linguistics (pp. 563-591): Springer.

Baayen, R. H., van Rij, J., de Cat, C., & Wood, S. N. (2018). Autocorrelated errors in experimental data in the language sciences: Some solutions offered by Generalized Additive Mixed Models. In D. Speelman, K. Heylen & D. Geeraerts (Eds.), Mixed effects regression models in linguistics (pp. 49-69). Berlin: Springer.

Baayen, R. H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, 94, 206-234.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. Journal of Memory and Language, 59, 457-474.

Connell, K., Hüls, S., Martinez-Carzia, M., Qin, Z., Shin, S., Yan, H., & Tremblay, A. (2018). English learners' use of segmental and suprasegmental cues to stress in lexical access: An eye-tracking study. *Language Learning*, 68, 635-668.

References

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137, 151-171.

Jesse, A., Poellmann, K., & Kong, Y. Y. (2017). English listeners use suprasegmental cues to lexical stress early during spoken-word recognition. *Journal of Speech, Language, and Hearing Research*, 60, 190-198.

McQueen, J. M., & Viebahn, M. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, 60, 661-671.

Nixon, J. S., van Rij, J., Mok, P., Baayen, R. H., & Chen, Y. (2016). The temporal dynamics of perceptual uncertainty: Eye movement evidence from Cantonese segment and tone perception. *Journal of Memory and Language*, 90, 103-125.

Porretta, V., Kyröläinen, A., van Rij, J., & Järvikivi, J. (2017). VWPre: Tools for preprocessing visual world data, R package version 1.0.1.

Porretta, V., Kyröläinen, A.-J., van Rij, J., & Järvikivi, J. (2018). Visual world paradigm data: From preprocessing to nonlinear time-course analysis. In I. Czarnowski, R. Howlett & L. Jain (Eds.), Intelligent Decision Technologies 2017, number 73 (pp. 268-277).

Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics*, 58, 1-21.

Reinisch, E., Jesse, A., & McQueen, J. M. (2010). Early use of phonetic information in spoken word recognition: Lexical stress drives eye movements immediately. *Quarterly Journal of Experimental Psychology*, 63, 772-783.

References

Sóskuthy, M. (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics*, 84.

Sulpizio, S., & McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language*, 66, 177-193.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.

van Rij, J., Hendriks, P., an Rijn1, H., Baayen, R. H., & Wood, Simon N. (2019). Analyzing the time course of pupillometric data. *Trends in Hearing*, 23, 1-22.

van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, H. (2017). itsadug: Interpreting time series and autocorrelated data using GAMMs.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, 70, 86-116.

Wood, S. N. (2006). Generalized additive models: An introduction with R. Boca Raton [u.a.]: CRC Press.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73, 3-36.

Wood, S. N. (2017). Generalized additive models: An introduction with R (2nd ed.). Boca Raton [u.a.]: CRC press.

References

Zahner, K., Kutscheid, S., & Braun, B. (2019). Alignment of f0 peak in different pitch accent types affects perception of metrical stress. *Journal of Phonetics*, 74, 75-95.

Zahner- Ritter, K., Chen, Y., Dehé, N., & Braun, B. (Accepted). The prosodic marking of rhetorical questions in Standard Chinese *Journal of Phonetics*.

Zahner-Ritter, K., Einfeldt, M., Wochner, D., James, A., Dehé, N., & Braun, B. (2022). Three kinds of rising-falling contours in German wh-questions: Evidence from form and function. *Frontiers in Communication*. https://doi.org/10.3389/fcomm.2022.838955

Zahner-Ritter, K., Kutscheid, S., & Braun, B. (2022). How experience with high and low pitch accents affects the cue weights in stress processing: Evidence from exposure-test paradigms using eye-tracking. Talk at the 55th Annual Meeting of the Societas Linguistica Europaea. Bucharest, Romania.

Zahner-Ritter, K., Zhao, T., Einfeldt, M., & Braun, B. (2022). How experience with tone in the native language affects the L2 acquisition of pitch accents. Frontiers in Psychology, https://doi.org/10.3389/fpsyg.2022.903879.