Variability in eye movements during reading: Bayesian inference of the SWIFT model

Maximilian M. Rabe, Johan Chandra, André Krügel, Stefan A. Seelig, & Ralf Engbert

Research Focus Cognitive Sciences, University of Potsdam, Germany

Universität Potsdam

SFB 1287 Limits of Variability in Language: Cognitive, Grammatical, and Social Aspects

Gefördert durch DFG Deutsche Forschungsgemeinschaft

https://doi.org/10.17605/osf.io/s5jk2

Introduction

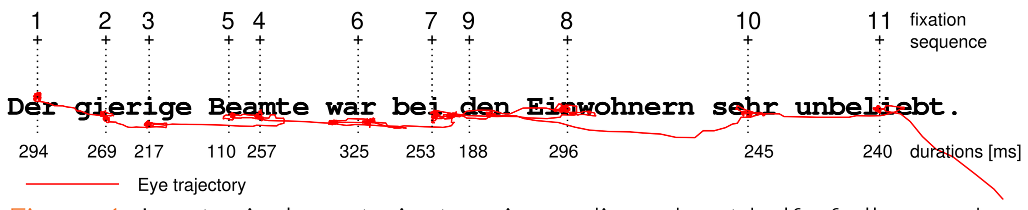

Figure 1. In a typical eye trajectory in reading, about half of all saccades move the foveal part of the visual field from word n to the next word $n+1$. Other saccade types generate refixations (e.g. 2–3), word skipping (e.g. 7–8), and regressions (8–9).

SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) is a dynamical cognitive model of eye-movement control in reading. Words within the processing span centered around the current fixation position are processed in parallel via a temporally evolving activation field (see Figure 2).

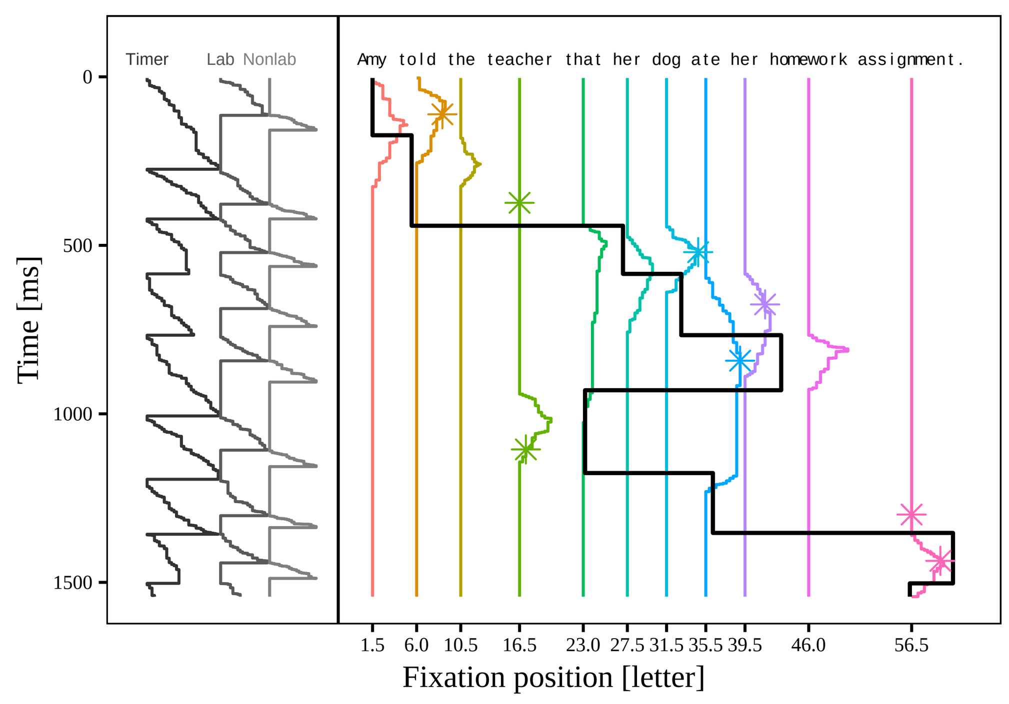

Figure 2. Simulated eye trajectory (solid black line) in the SWIFT model. Thin lines are word activations (colored) and timers (gray) as a function of time. Asterisks mark points in time when a saccade program is executed.

Recently, we implemented SWIFT for Bayesian parameter estimation (Seelig et al., 2019). We fitted the model to a diverse reading dataset in order to evaluate the goodness of fit with various temporal and spatial summary statistics.

For the simulations reported here, we replaced the standard Gaussian saccade error model with Gamma-distributed saccade amplitudes.

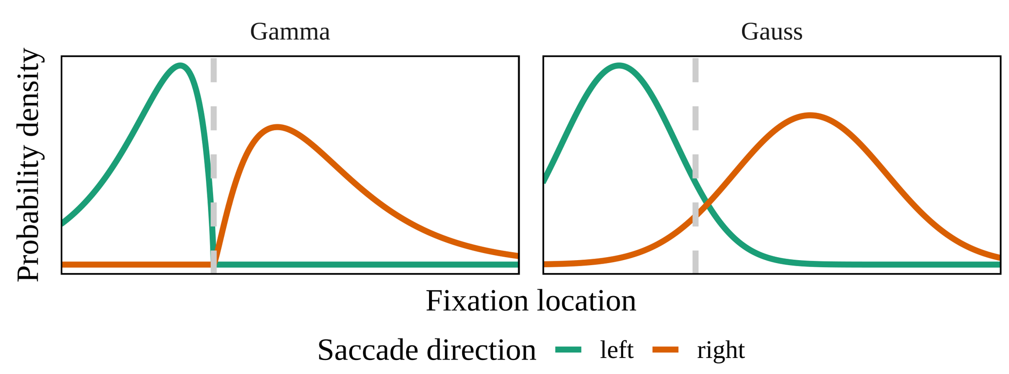

Figure 3. Comparison of saccade amplitude distributions for saccade targets left or right to the current fixation location (dashed gray line).

Bayesian parameter estimation

Bayesian model fitting allows us to infer rigorous credibility intervals for model parameters. Additionally, we can determine model parameters *for individual participants*, which was often precluded in previous methods.

In SWIFT, the likelihood of a fixation $f_i = (k_i, l_i, T_i)$ is given as the combined spatial and temporal likelihoods, both conditional on all preceding fixations $F_{i-1} = (f_2, ..., f_{i-1})$, model parameters θ and internal degrees of freedom ξ:

$$L_M(k_i, l_i, T_i | F_{i-1}, \theta, \xi) = L_{temp}(T_i | F_{i-1}, \theta, \xi) \cdot L_{spat}(k_i, l_i | T_i, F_{i-1}, \theta, \xi)$$

Likelihood profiles for selected parameters show that for simulated data, (a) the true parameter values are most likely and (b) model parameters can have different selective influences on spatial and temporal likelihood components:

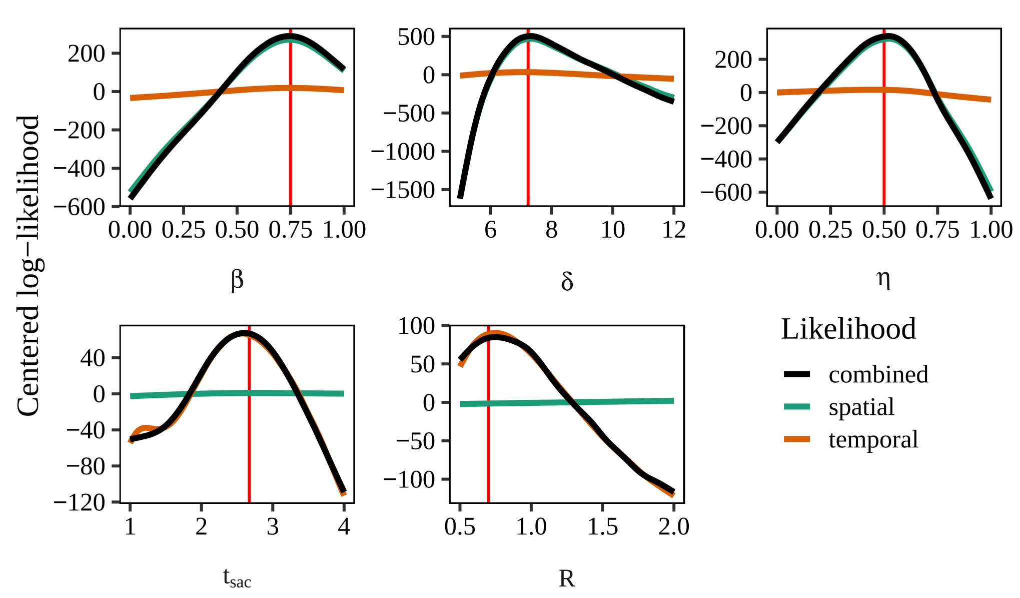

Figure 4. Centered log-likelihood profiles (black) with temporal and spatial components. Vertical red lines are true parameter values.

Computational modelling

- DREAMzs sampling algorithm (Vrugt et al., 2009)
- PyDREAM implementation (Shockley, 2019), modified to enable likelihood reevaluation
- Fitting to 70% of the data, posterior predictive checks (Figures 6–9) for the remaining 30%

Results

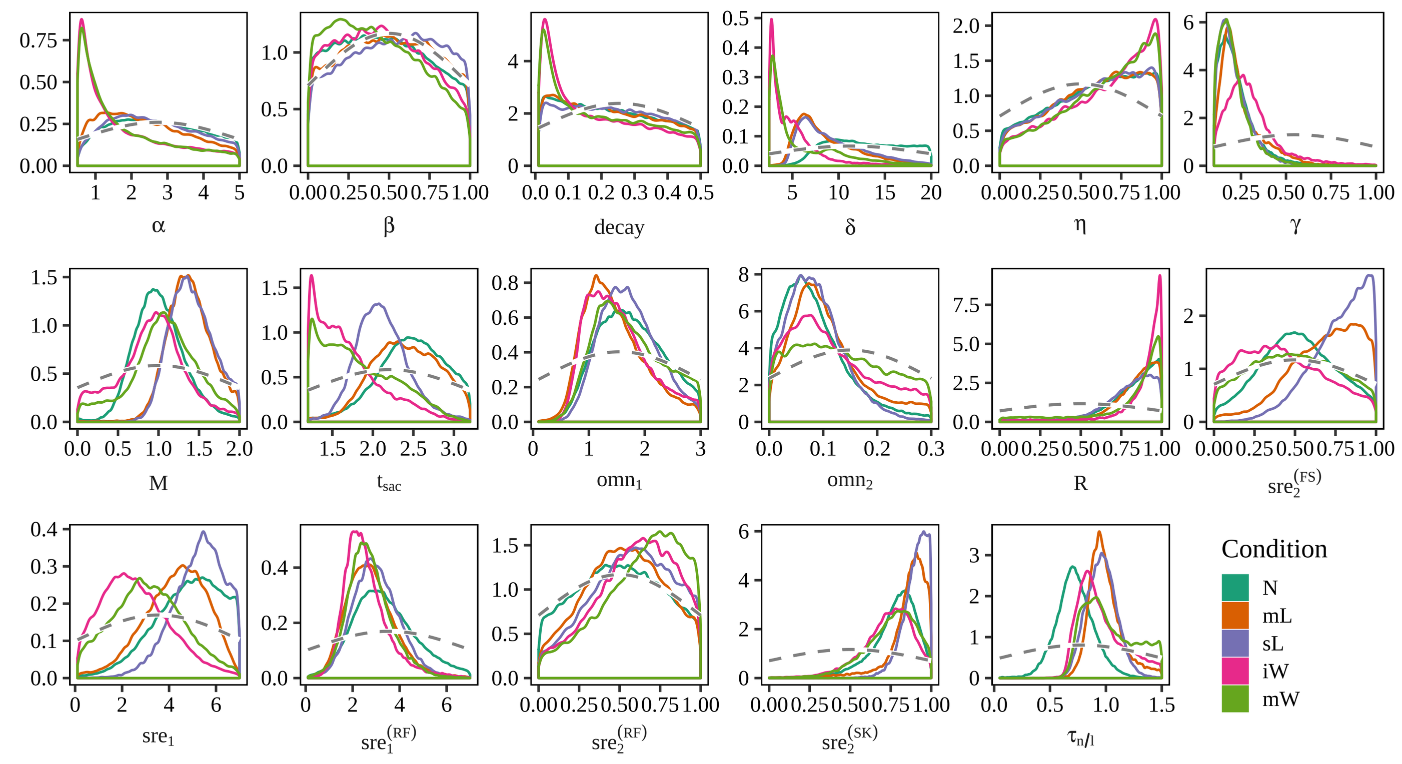

Figure 5. Posterior parameter distributions. Each color represents the aggregated sampled posteriors across all subjects in that condition. Priors (gray dashed lines) are truncated normal with support on 1 SD around the mean and were identical across subjects and conditions.

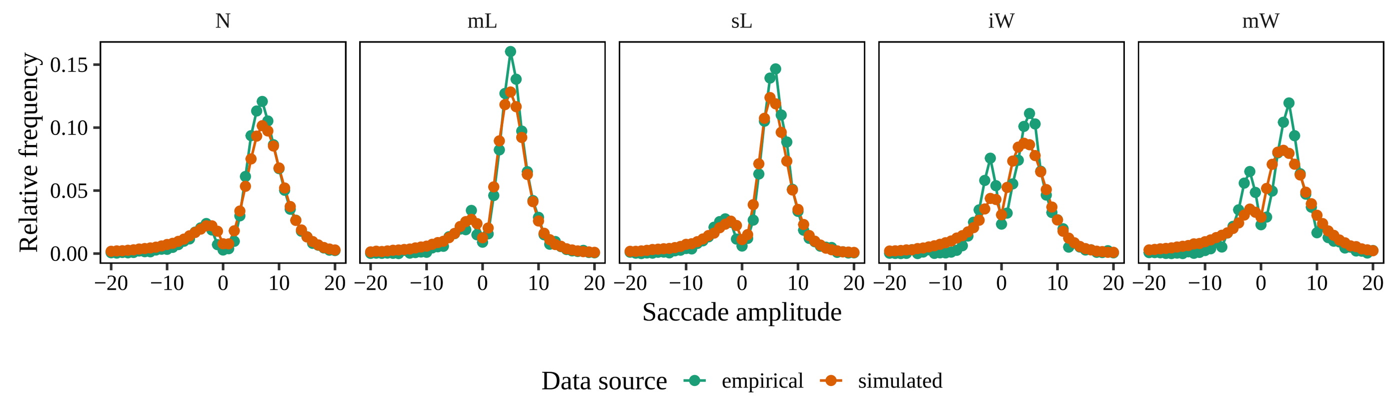

Figure 6. Empirical and simulated saccade amplitudes aggregated across all participants in each experimental condition, including the baseline condition (N).

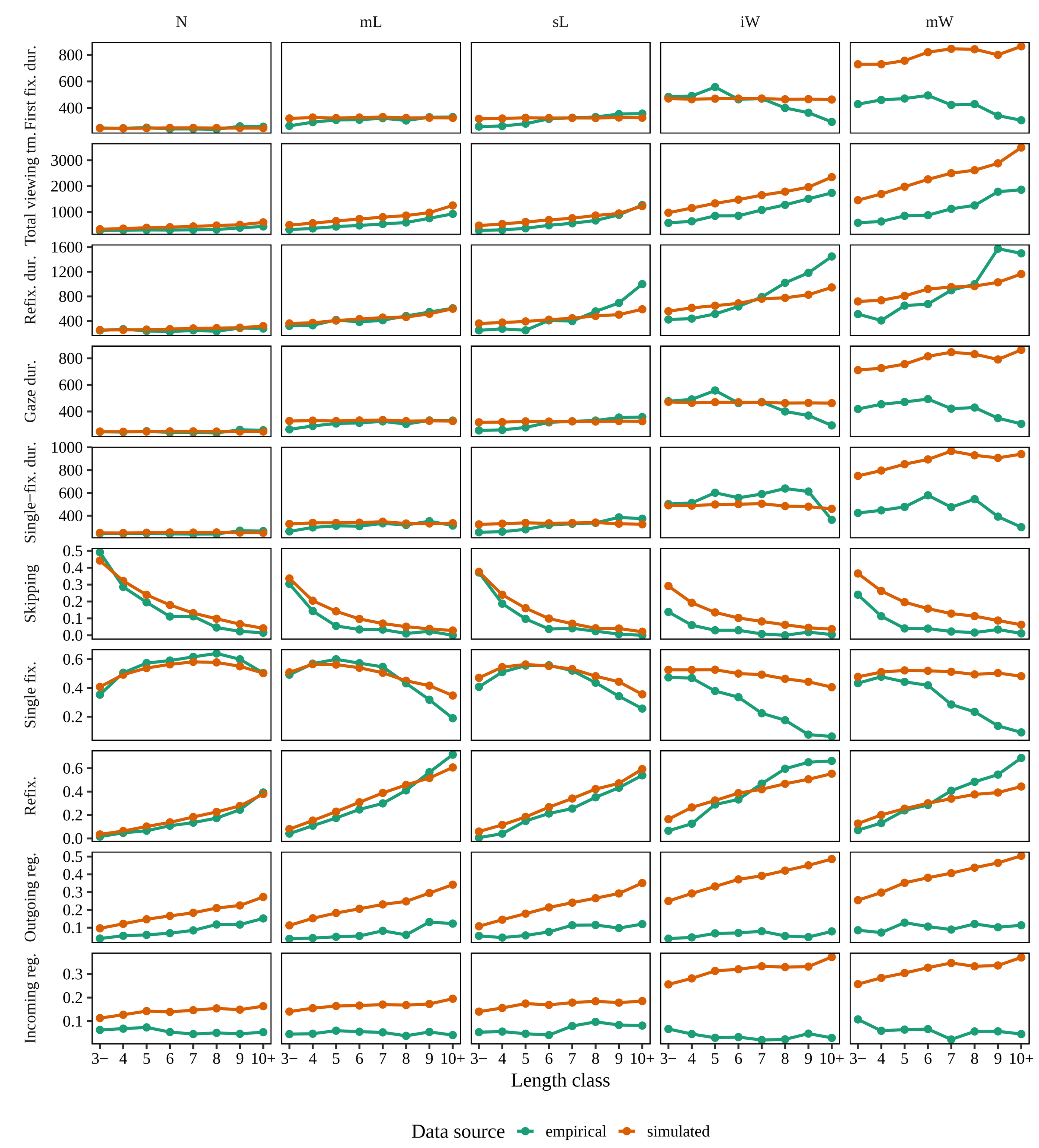

Figure 7. Empirical and simulated temporal (1–5) and spatial (6–10) summary statistics for different experimental conditions, aggregated across subjects, as a function of word length.

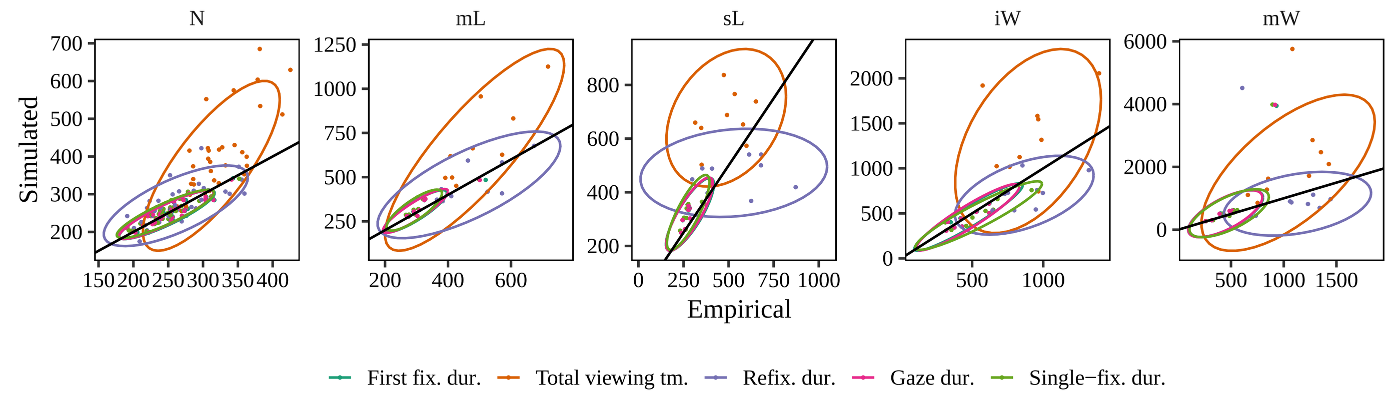

Figure 8. Correlation between empirical and simulated temporal summary statistics. Each participant is represented by one dot in each color in the respective experimental condition (panel).

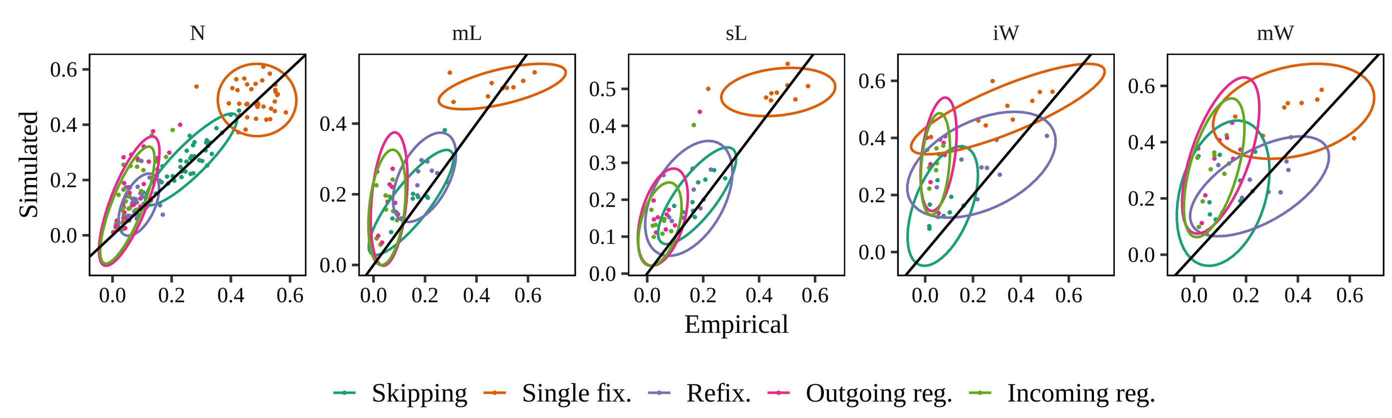

Figure 9. Correlation between empirical and simulated spatial summary statistics. Each participant is represented by one dot in each color in the respective experimental condition (panel).

Experimental method

Table 1. Reading conditions

N	Jede	Sprache	der	Welt	besitzt	eine	Grammatik
mL	Jede	Sprache	der	Welt	besitzt	eine	Grammatik
sL	Jdee	Scrahpe	der	Wlet	bsizett	enie	Gmartimak
iW	edeJ	ehcarpS	red	tleW	tztiseb	enie	kitammarG
mW	Jede	Sprache	der	Welt	besitzt	eine	Grammatik

- 36 native German speakers with normal or corrected-to-normal vision
- Normal reading (N) in first lab session
- One of four manipulated reading conditions (see Table 1) in second lab session

Parameter estimates

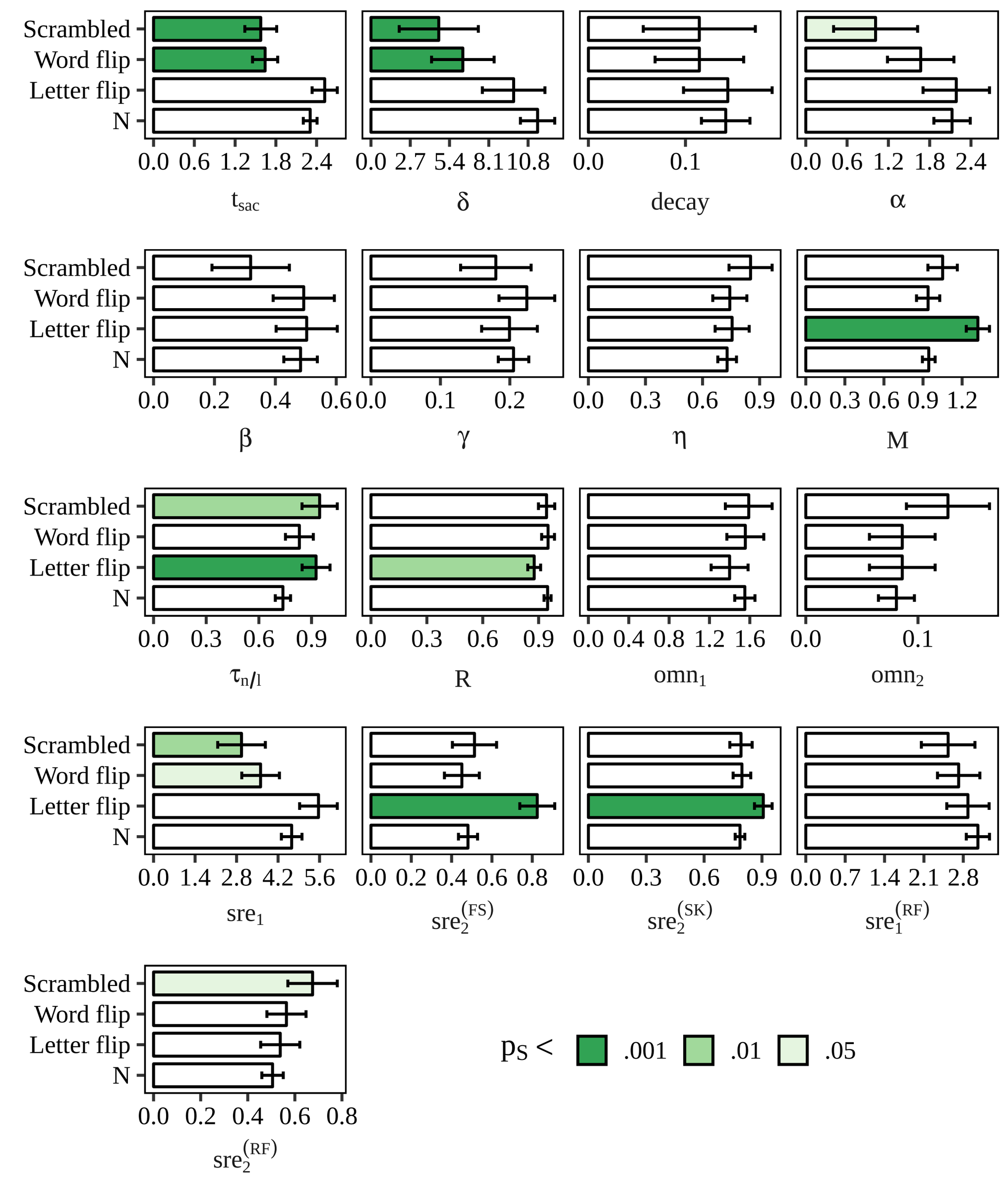

Figure 10. Linear regression coefficient estimates of model parameter estimates between experimental conditions. Error bars are 95% confidence intervals around the estimated means. The baseline is tested against zero, while conditions are tested against the baseline. p-values are corrected for independent multiple testing according to Šidák (1967) with $p_S = 1 - (1-p)^{17}$.

Reversing letter order within the word narrows the processing span, resulting in shorter and more frequent saccade programs.

Flipping letters horizontally (1) leads to more misplaced fixations due to greater systematic oculomotor error, (2) prolongs saccade programs following misplaced fixations and (3) increases the duration of the labile and non-labile saccade programs.

Scrambling letters within words slows down overall processing and is furthermore associated with (almost) the same patterns as flipping letters and words.

Summary

- SWIFT was successfully fitted to empirical data collected under different reading conditions
- Goodness of fit was ensured by comparing empirical and simulated summary statistics
- Subject-level parameters can reliably predict characteristics of reading patterns in unseen trials
- Differences in subject-level parameters could explain why and how differences in reading behavior arise

References

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813. doi: 10.1037/0033-295X.112.4.777

Seelig, S. A., Rabe, M. M., Malem-Shinitski, N., Risse, S., Reich, S., & Engbert, R. (2019). Bayesian parameter estimation for the SWIFT model of eye-movement control during reading. *Journal of Mathematical Psychology*. (Manuscript under revision)

Shockley, E. (2019). *PyDREAM*. Retrieved from `https://github.com/LoLab-VU/PyDREAM`

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, *62*, 626–633. doi: 10.1080/01621459.1967.10482935

Vrugt, J. A., ter Braak, C., Diks, C., Robinson, B. A., Hyman, J. M., & Higdon, D. (2009). Accelerating markov chain monte carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, *10*(3). doi: 10.1515/ijnsns.2009.10.3.273

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project no. 317633480 – SFB 1287, project B03. We acknowledge the North-German Supercomputing Alliance (HLRN) for providing high-performance computing resources that have contributed to the research results reported in this paper.