A Computational Model of Eye Movements During Reading

Reinhold Kliegl & Ralf Engbert
University of Potsdam

1. A few basic facts about eye movements in reading
2. Modeling the basic facts with SWIFT
3. Application to individual differences

Eye movements and fixations during reading

In der Klosterschule herrschen Schwester Agathe und Schwester Maria.
232

Es sollte mehr Strom mit Solarenergie erzeugt werden.
280

Viele Kinder lesen nur noch selten.
392

Gute Beziehungen ebnen vielen Unternehmern den Weg zum Erfolg.
380

Das Fenster im Flur klemmt seit ein paar Tagen.
360

Sarah hat ihrem Opa ein Bild gemalt.
376

Jede Sprache der Welt besitzt eine Grammatik.
344

Als Kapitalanlage ist Gold nicht zu empfehlen.
408

Eye Movements and Fixations During Reading

Duration

fixation: 30—750 ms

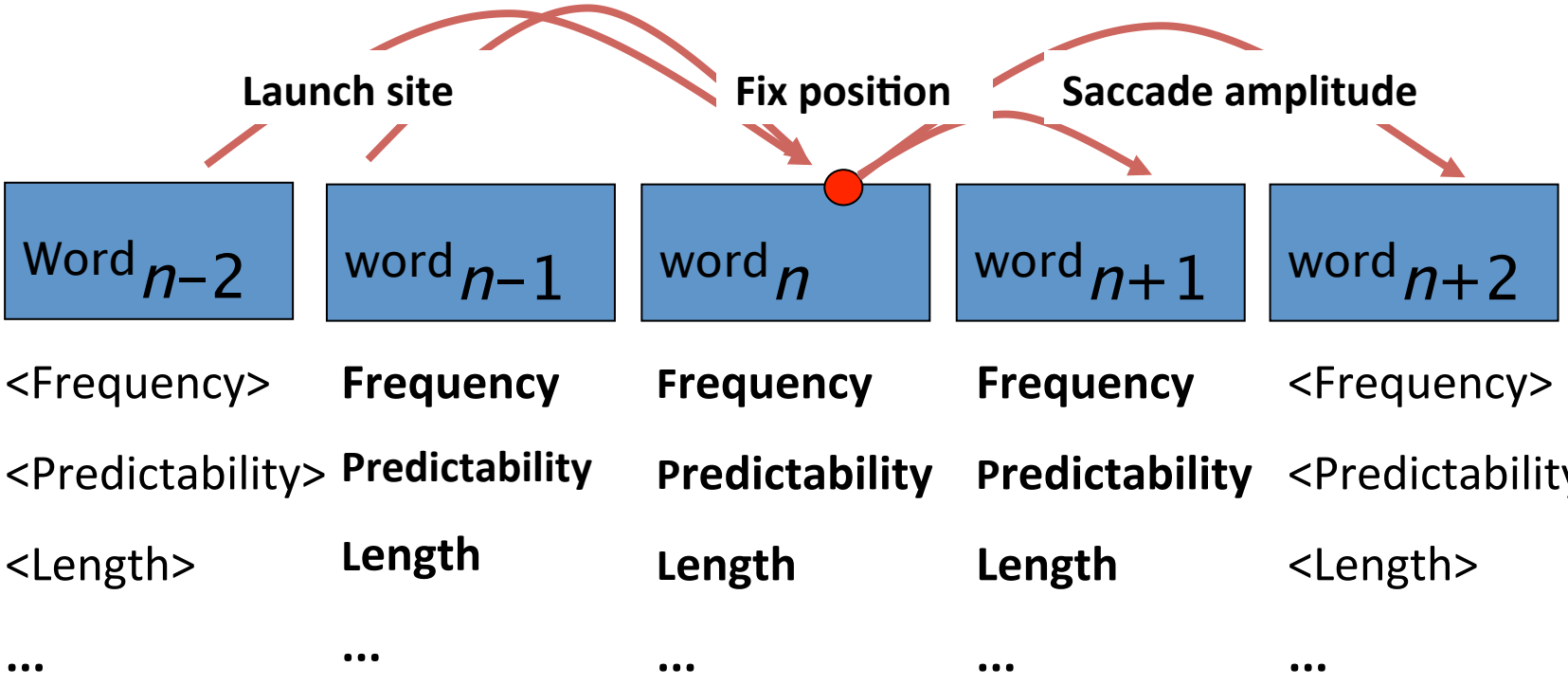
saccade: 20—40 ms

Data

80 530 single fixation durations (1st-pass)

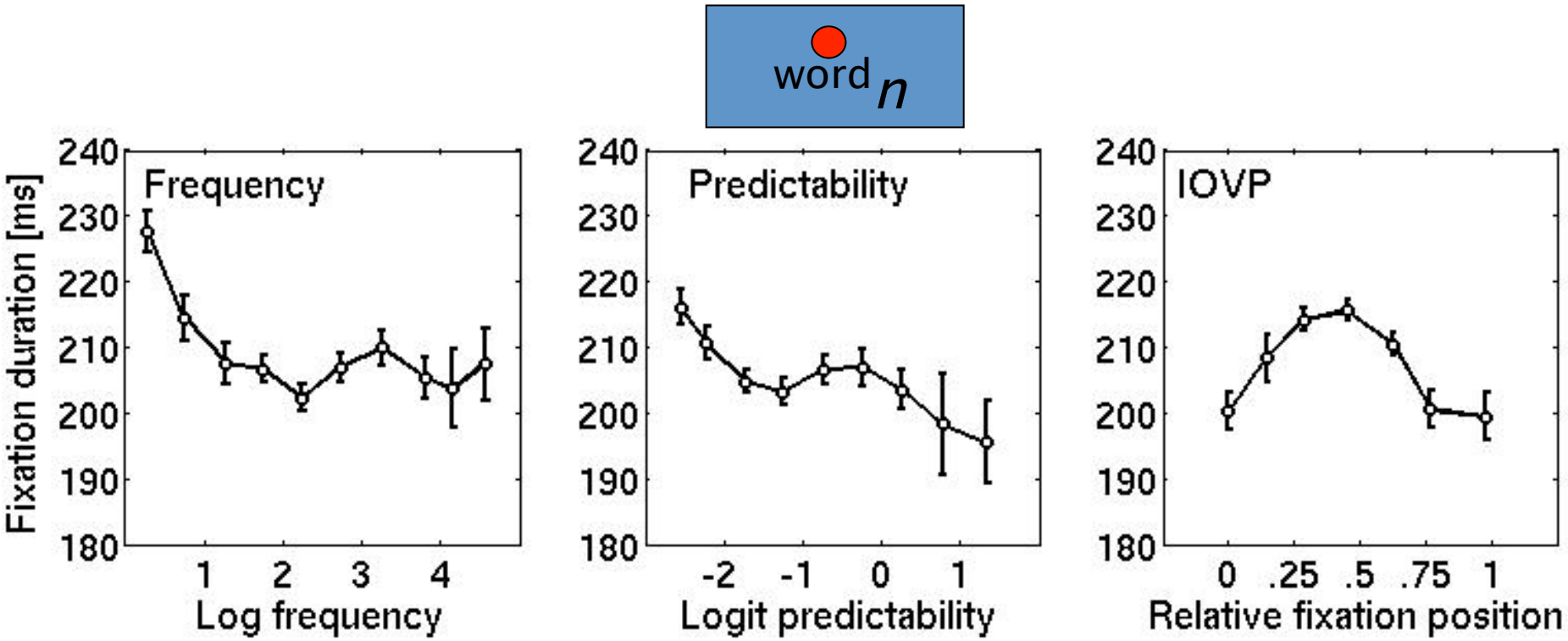
Random factors in linear mixed model:

subject (273); sentence (144); word (369)



Kliegl, 2007; Heister, Würzner, & Kliegl (2012)

Immediate Effects of Frequency, Predictability, & Fixation Position: Potsdam Sentence Corpus

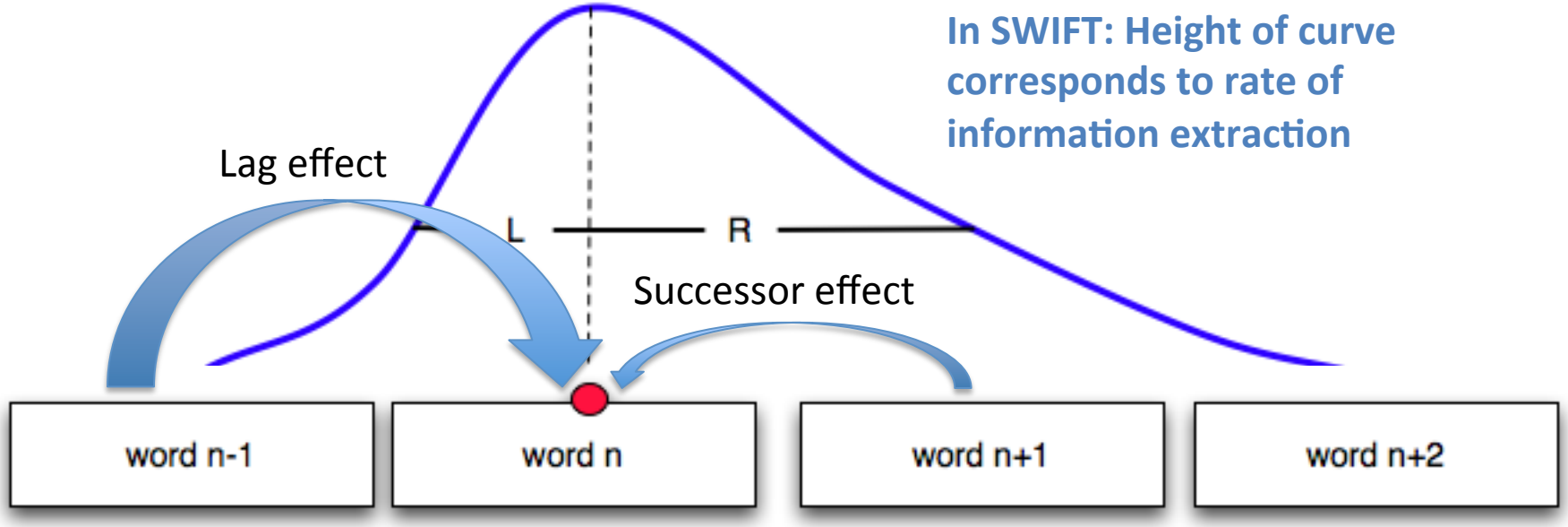


Predictability = prob(guessing word n | words 1 to n-1); 83 guesses/wrd;

metric: logits [log of odds; e.g., log(1:1) = 0].

Kliegl, Nuthmann, & Engbert (2006)

Distributed Processing Effects During Reading



Launch site

Frequency
Predictability
Semantics

…

Position

Frequency
Predictability
Semantics

…

Orthography
Phonology
Frequency?
Predictability?
Semantics?
Morphology?

…

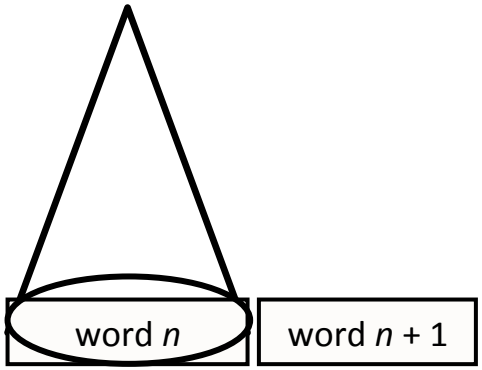
Orthography?

…

Models of Eye Movement Control: Visual Attention

Attention Spotlight

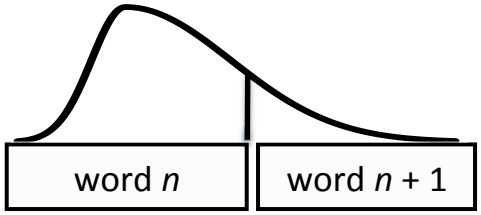
e.g.,
- Morrison
- E-Z Reader



Reichle et al. (1998, 2008)

Attention Gradient

e.g.,
- SWIFT 2
- Glenmore

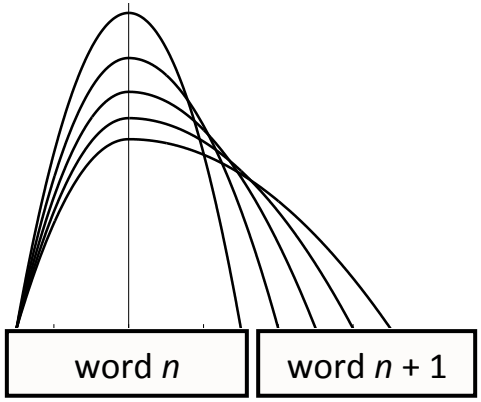


Engbert et al. (2002, **2005**)

Zoom-lens of Attention

Eriksen & St James, 1986

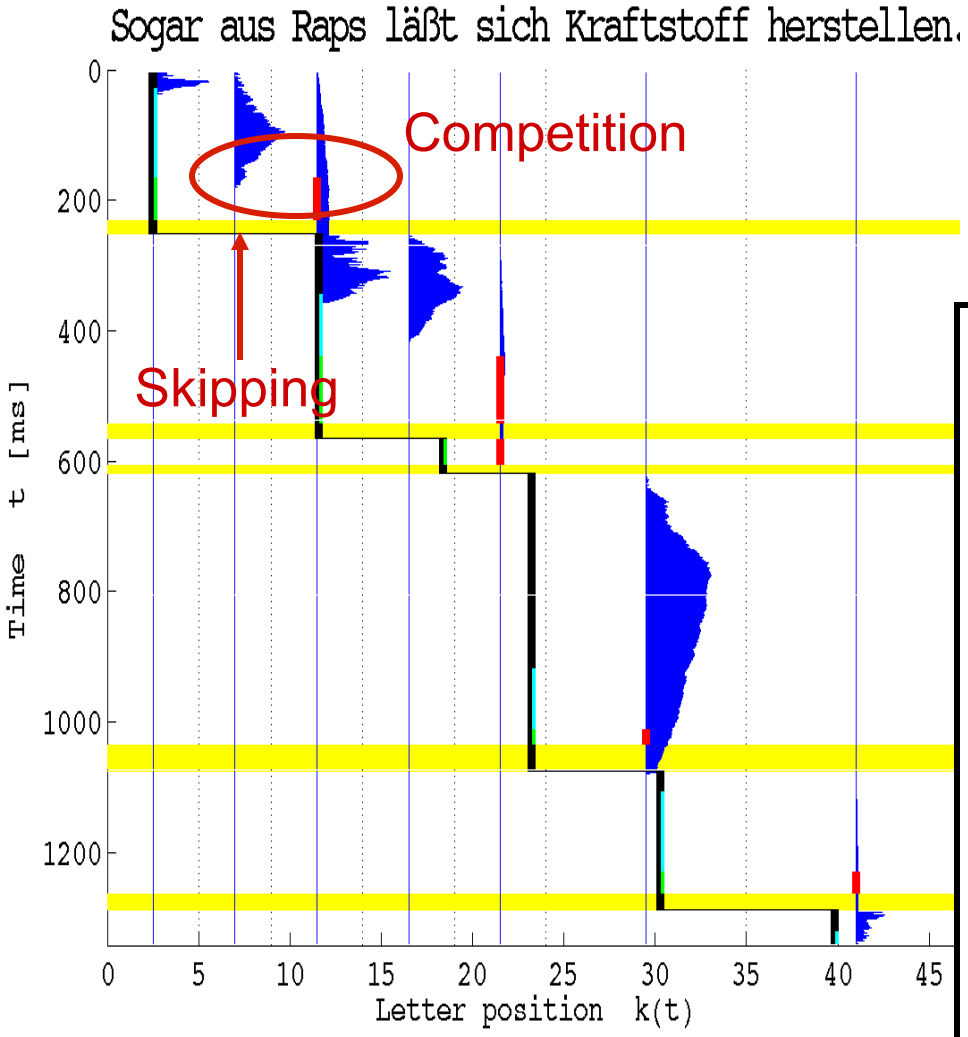
→ SWIFT 3



Schad & Engbert (2012)
Risse, Hohenstein, Kliegl, & Engbert, 2014)

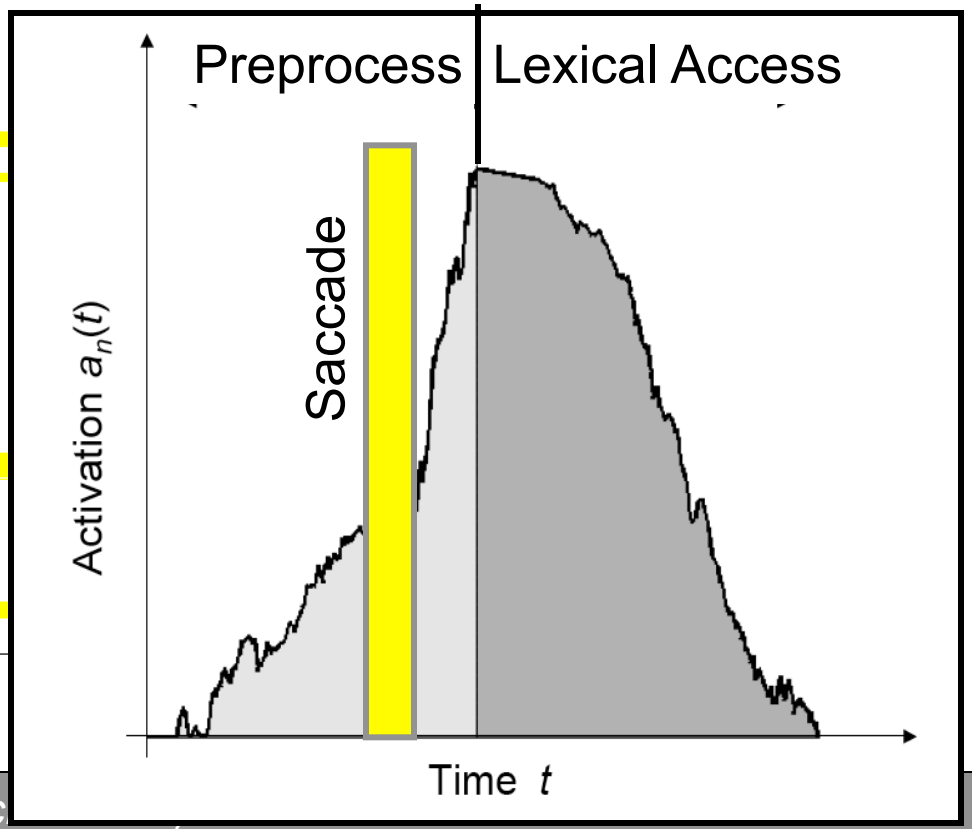
Distributed processing during eye fixations in reading

Example of SWIFT Simulation



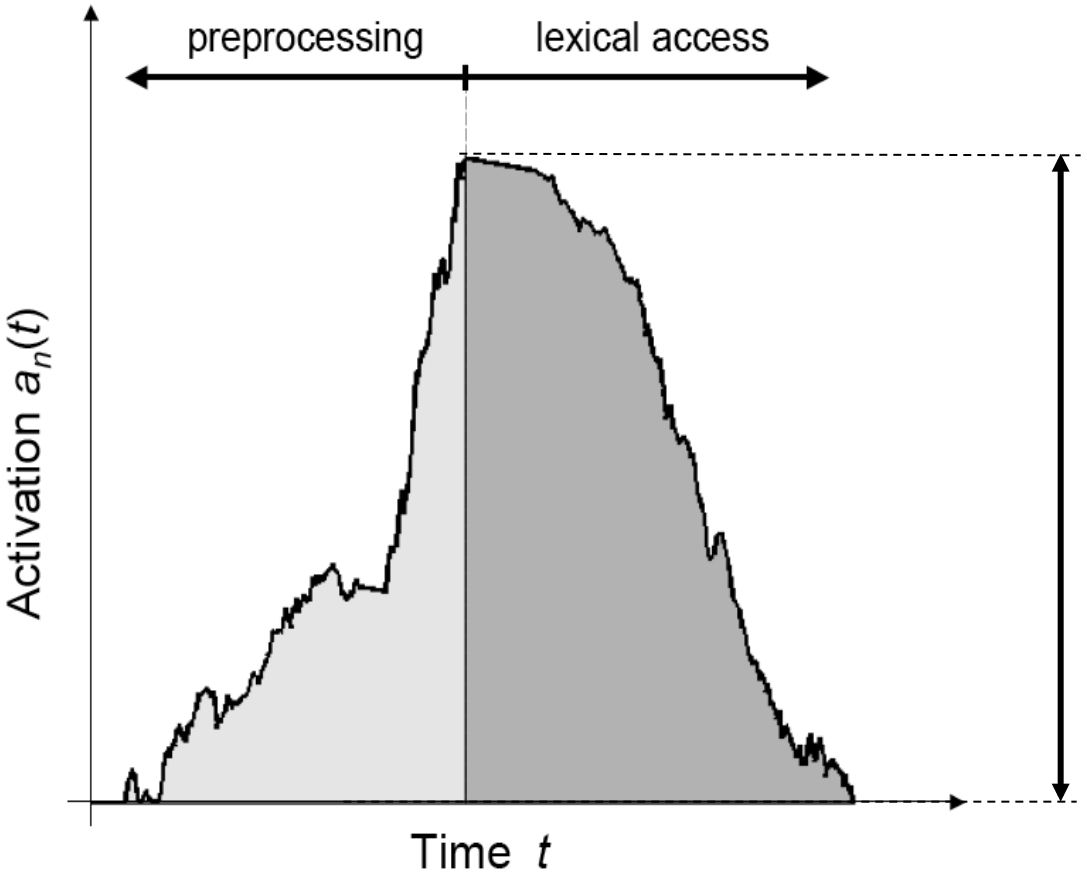
Theoretical principles

1. Spatially distributed processing of words („activation field“)
2. Separate pathways for saccade target selection and timing



Engbert, Nuthmann, Richter, & Kliegl (2005, *Psyc*

Word difficulty



Lexical difficulty of word *n*

$$L_n = \alpha - \beta \log f_n$$

f_n: printed word frequency

Engbert *et al.* (2005) *Psychol. Rev.*

Saccade target selection: A stochastic process

Probability computed from dynamic field

$$\pi(n,t) = \frac{a_n^{\gamma}(t)}{\sum_{j=1}^{N} a_j^{\gamma}(t)}$$

$\gamma = 0$: random target selection

$\gamma \to \infty$: winner-takes-all

Result from numerical simulations:
(Luce's unbiased choice rule, 1959)

$\gamma \approx 1$

Temporal evolution of activations

Equation of motion: *N* coupled differential equations

$$\frac{\mathrm{d}a_n}{\mathrm{d}t} = F_n(t)\Lambda_n(t) - \omega$$

$F_n(t)$ preprocessing factor, modulated by word predictability

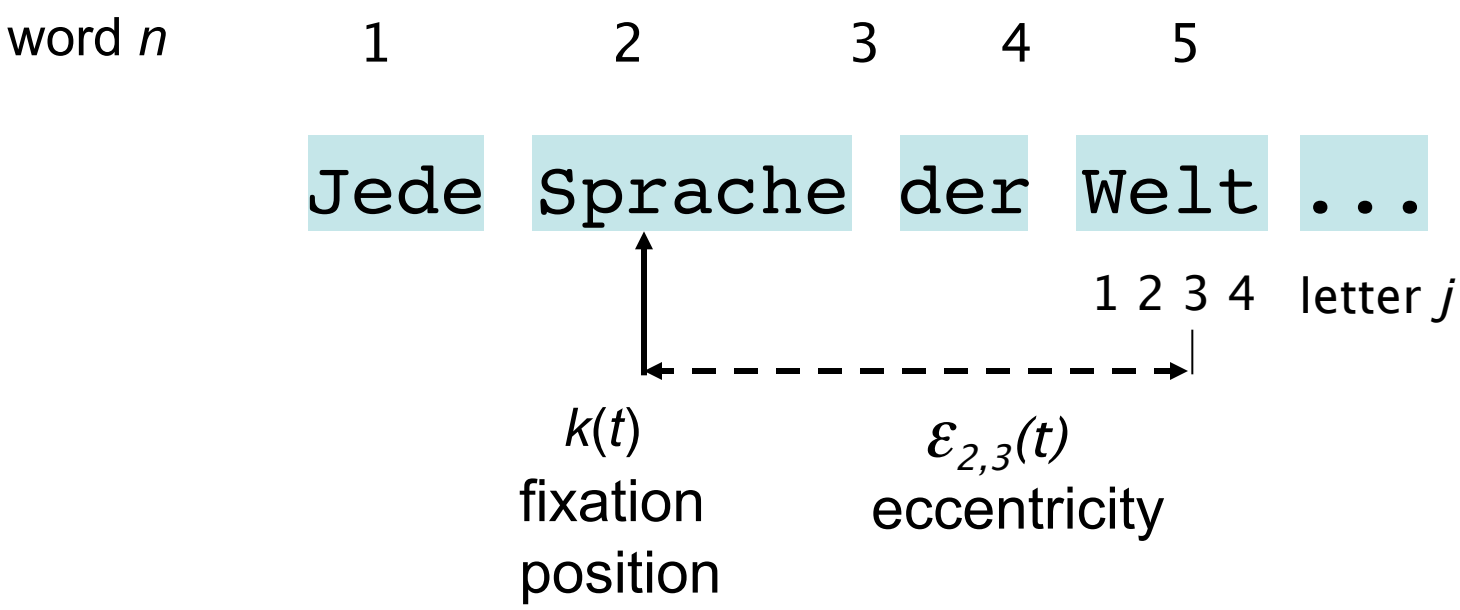
$\Lambda_n(t)$ stochastic processing rate: random walk (e.g., Ratcliff, 1978)

ω global (constant) decay of activation

From perceptual span to processing rates

Eccentricity of letter *j* of word *n*

$$\epsilon_{nj}(t) = x_{nj} - k(t)$$



From perceptual span to processing rates (cont'd)

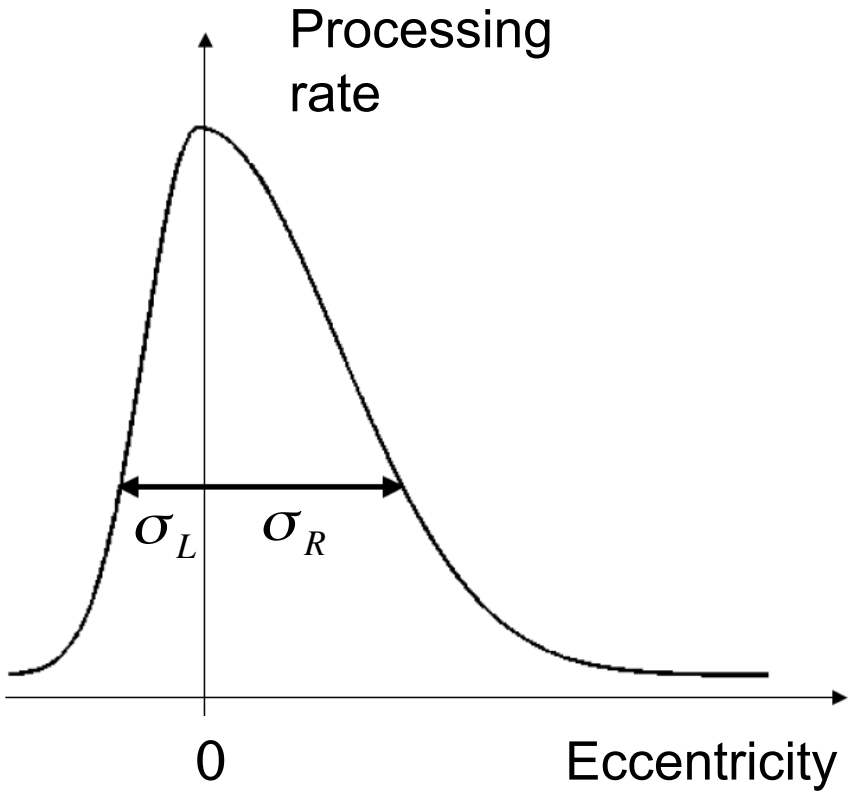
Eccentricity $\varepsilon_{nj}(t)$ of letter *j* of word *n*

↓

Gaussian-distributed processing rates for letters $\lambda(\varepsilon_{nj}(t))$

↓

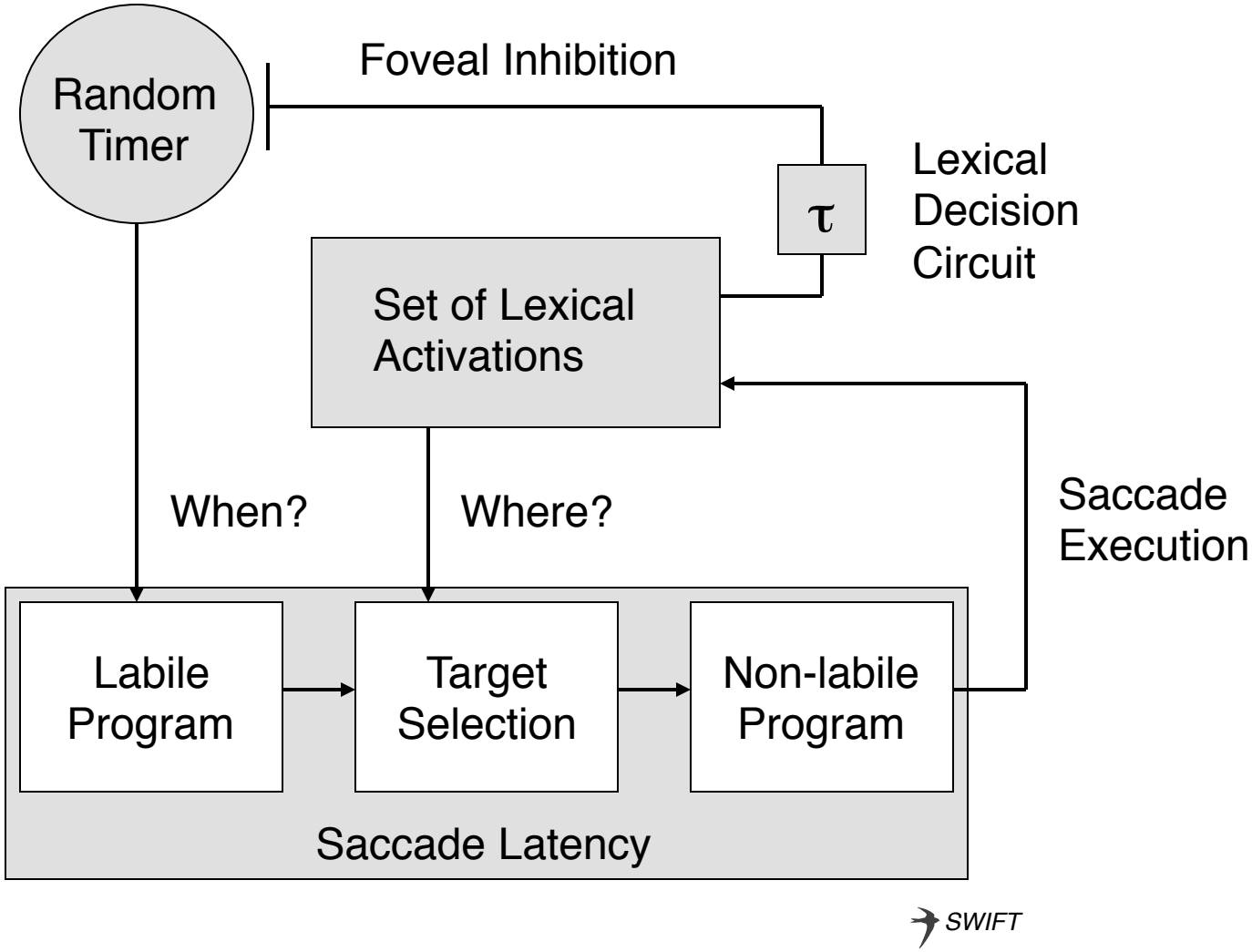
Processing rate of word *n*



$$\lambda_n(t) = \frac{1}{(M_n)^\eta} \sum_{j=1}^{M_n} \lambda(\epsilon_{nj}(t))$$

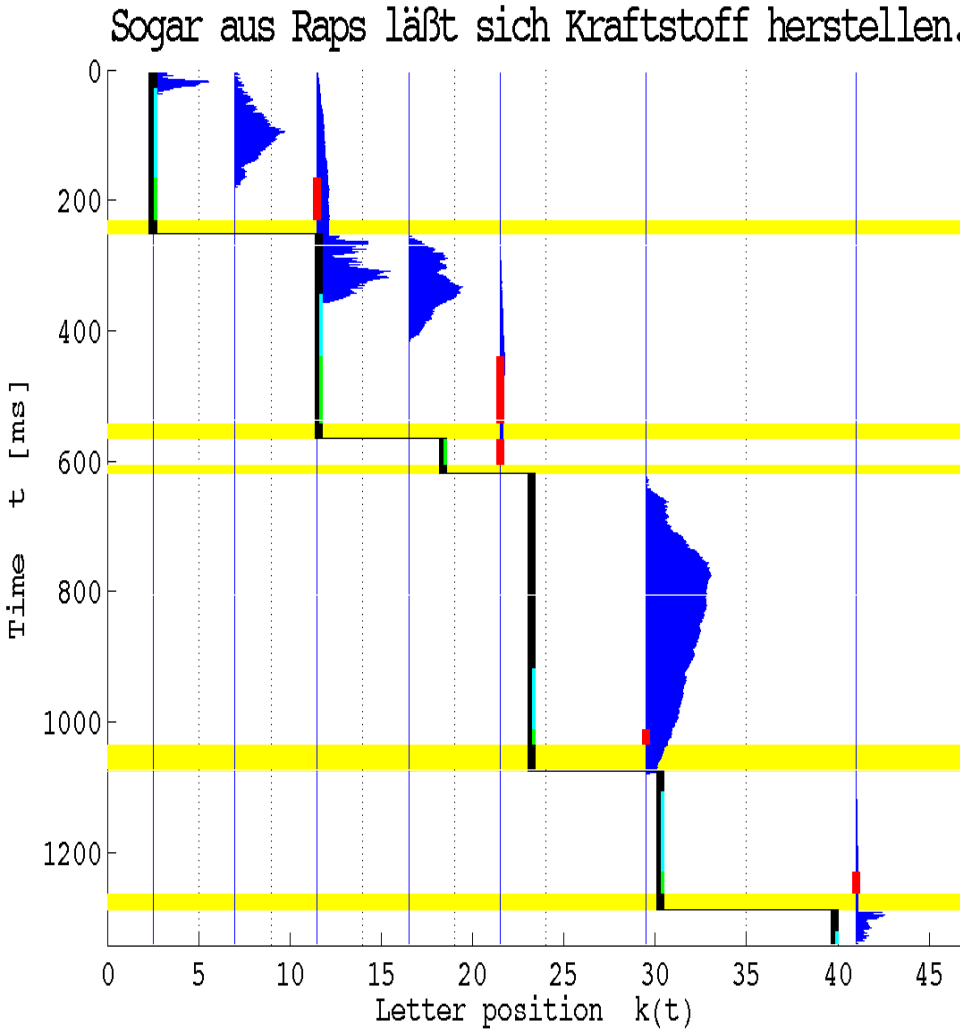
Engbert et al. (2005) *Psychol. Rev.*

The SWIFT model: An overview



SWIFT

Example of SWIFT Simulation



Theoretical principles

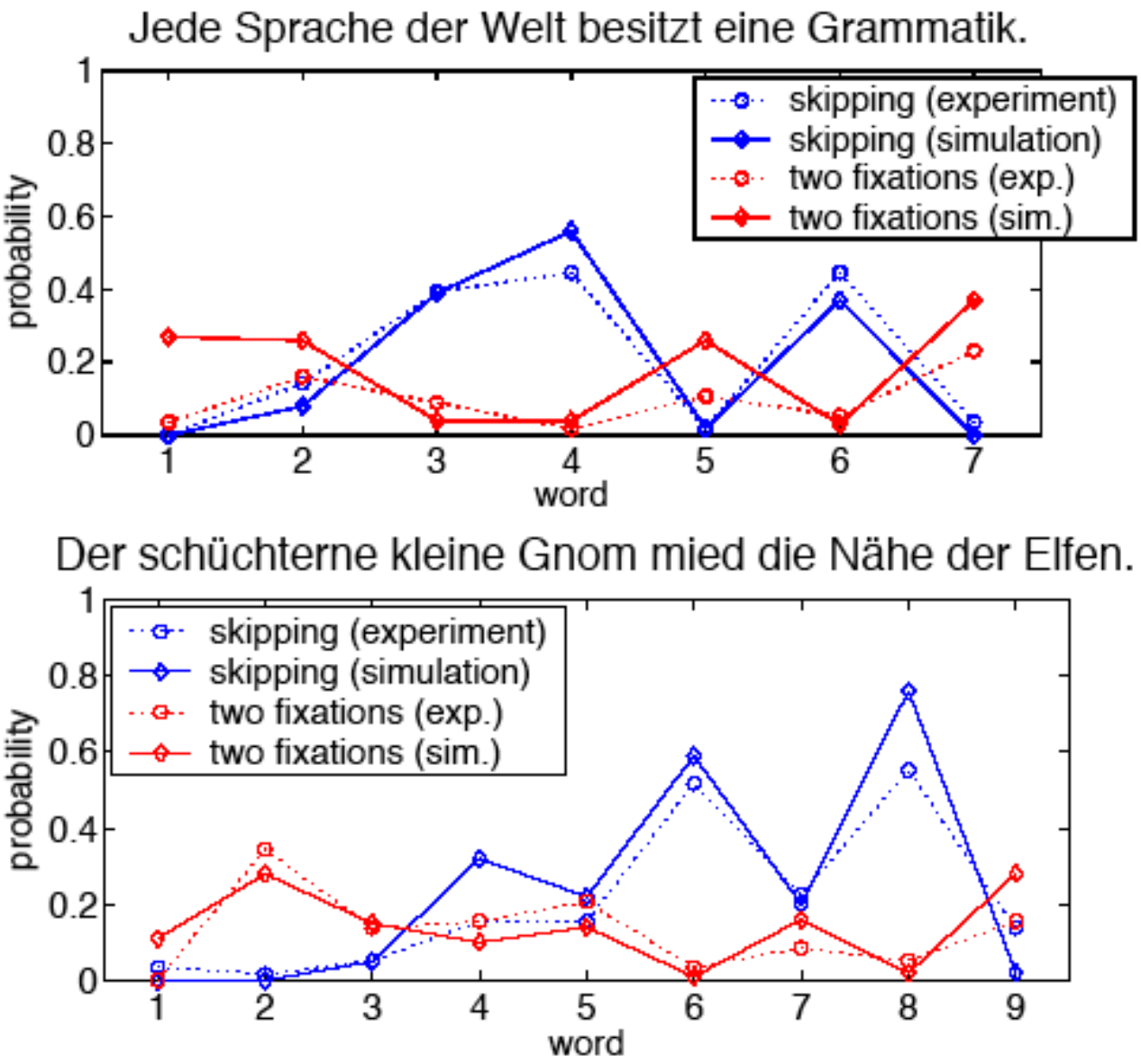
1. Parallel processing of words in the perceptual span
2. Target selection according to lexical activation of „field“ of words
3. Autonomous saccade programs, modulated by foveal inhibition

4. - 7. Sacc. program, oculomotor control

Model fitting

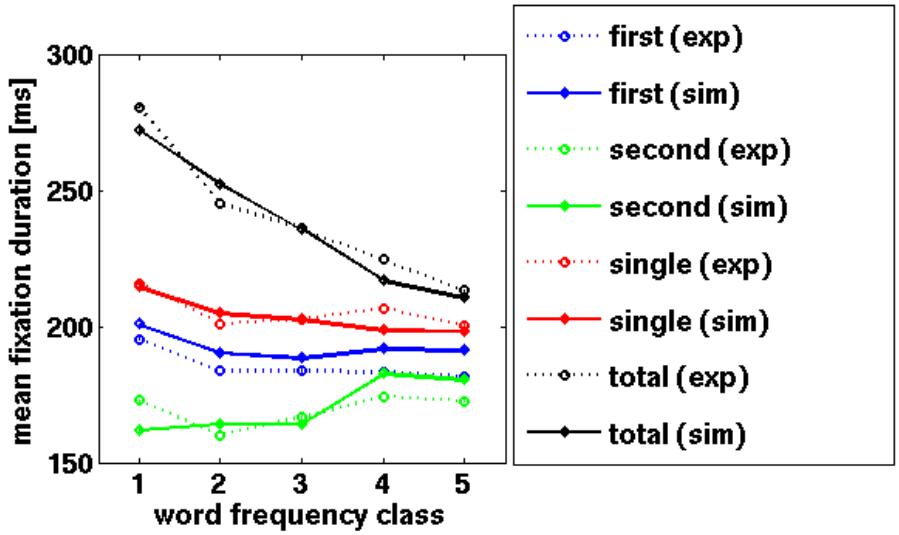
1. 222 readers, ~ 220 K fixations
2. For each of 850 words: Mean of 4 fix probs (0, 2, 3+, regression), 4 fix durs (single, 1st, 2nd, total), +
3. ~ 6800 means, 13 free parameters, estimated w/ genetic algorithm

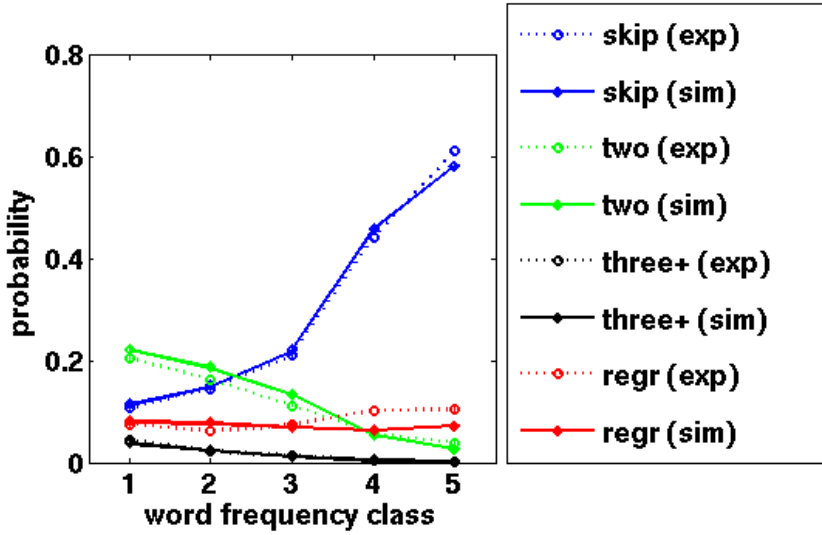
Numerical Results: Word-based Measures

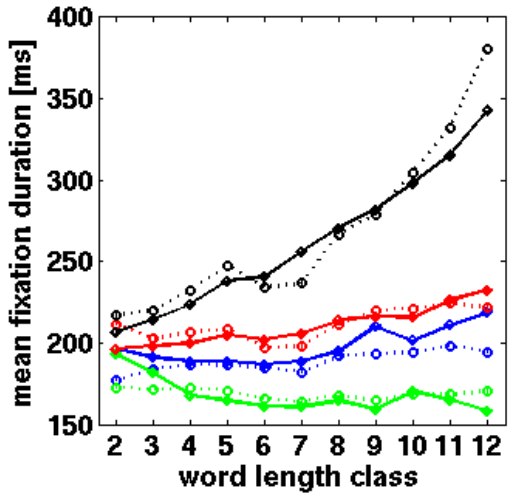


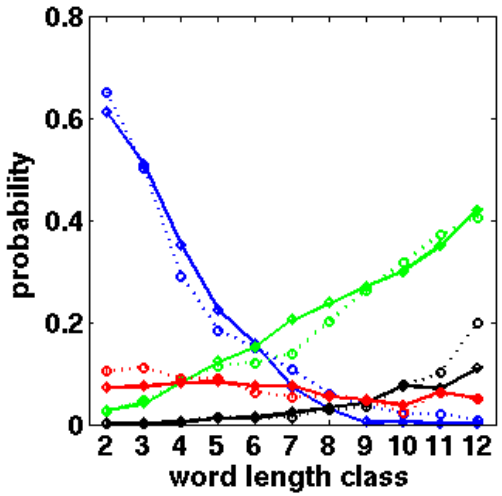
Engbert et al. (2005, *Psychol. Rev.*)

Recovery of Fixation Durations and Probabilities



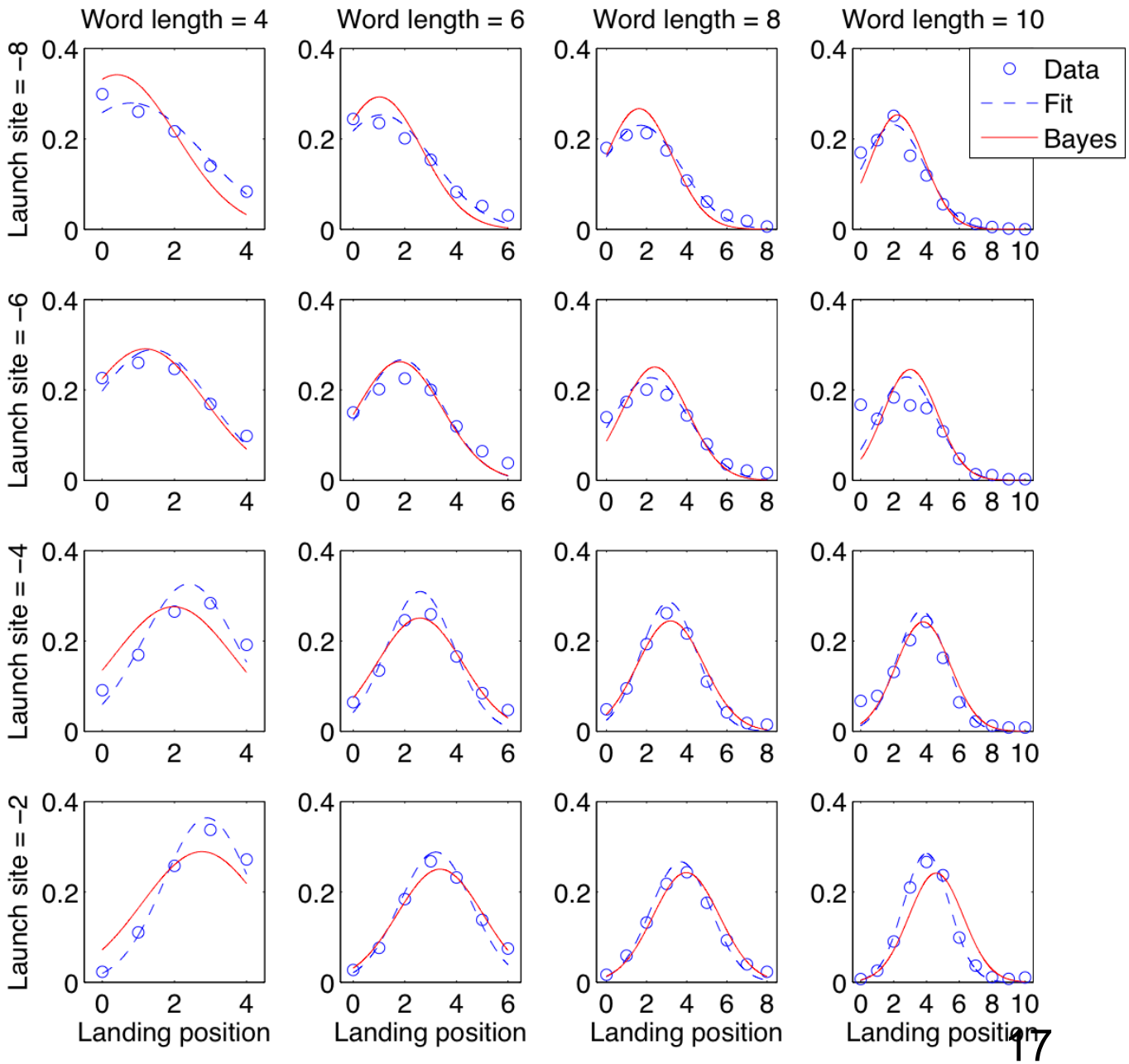






Engbert et al. (2005, *Psychol Rev*)

Results: Landing-Position Distributions



Thank you

Ralf Engbert
Website: http://www.agnld.uni-potsdam.de/~ralf/

Supported by

Deutsche Forschungsgemeinschaft DFG

PP-Design:
Laura Risse