Distributed Processing Across Reading Fixations

- word n-2: frequency, predictability, saccade amplitude
- word n-1: frequency, predictability
- word n: (highlighted)
- word n+1: frequency, predictability
- word n+2: frequency

**Dep. variable**: first-pass single-fixation durations (over 90 000)

**Word frequency**: DWDS corpus (125 million words)

**Word predictability**: prob(guessing word n | words 1 to n-1); 83 guesses/wrd

**Fixation data**: 222 subjects, collapsing across 9 experiments

**Analysis**: linear mixed-effect models, using *lmer* (*lme4* package in R)

**CONTEXT**
- experiments:
  - instruction, bitebar, ...
- quasi-experiments:
  - age, expertise, ...

**@ reader**
- instruction, bitebar, ...
- age, expertise, ...

**@ sentence/word**
- preview, contrast, ...
- grammar, lexical status, ...
Predictability of the next word in reading increases fixation duration on the earlier word (Method: multivariate statistics; Kliegl et al., 2006)
Successor Effect of Memory
Predictability (Inverted) of $n+1$

Predictability of word $n$

Predictability of word $n+1$

Graphs showing the relationship between fixation duration on word $n$ and $n+1$ with logit predictability.
Anticipatory retrieval of word n+1

- Some words can be predicted from sentence context
  - Need no (or only minimal) visual information for recognition
  - Tend to be skipped, approached with longer saccades (O'Regan, '79)

- Word n+1 partially processed during fixation on word n
- Predictability effect reduced, not eliminated if word n+1 is fixated

**Successor Effect of Memory Predictability (Inverted) of n+1**

**Predictability of word n+1**
Bilingual Reading of German and English Sentences

- 30 English-German readers; 30 German-English readers
- UMass Amherst Corpus (48 sentences; Schilling et al., 1998; Reichle et al., 1998)
- Potsdam Sentence Corpus (144 sentences; Kliegl et al., 2004)
Features of Chinese Script

• Example

医生提醒市民们傍晚乘凉时尽量少去草丛茂密的地方。

*The doctor reminded the citizens that they should avoid lush grass*

• Each character takes same amount of horizontal extent with different levels of visual complexity.

• The majority of linguistically defined Chinese words are one and two characters in length (Yu et al., 1985).

• A Chinese character typically occupies the space of 3 letters in alphabetic languages (i.e., Tsai & McCokie, 1995), but carries comparatively more information about meaning.
Reading Sentences in Simplified Chinese: Beijing Sentence Corpus

- 60 readers of Simplified Chinese (Beijing)
- 150 sentences
- 32,414 single-fixation durations (firstpass reading)

Kliegl, Yan, et al. (2012)
Reading Spanish Proverbs and High and Low Predictable Sentences: Max-Jump of Predictability

- 41 Argentinian readers of Spanish (Bahia Blanca & Buenos Aires)
- 184 sentences (75 low predictable, 45 high predictable, 64 proverbs)
- 19,550 single-fixation durations (firstpass reading)

Fernández, Shalom, Kliegl, & Sigman (2014)
Reading Spanish Proverbs and High and Low Predictable Sentences: Before and After Pred-Jump

41 Argentinian readers of Spanish (Bahia Blanca & Buenos Aires)

184 sentences (75 low predictable, 45 high predictable, 64 proverbs)

19,550 single-fixation durations (firstpass reading)

Low predictable
High predictable
Proverb

Fernández, Shalom, Kliegl, & Sigman (2014)
II. Age of reader and visibility of text

Reinhold Kliegl, Antje Nuthmann, Jochen Laubrock, Sarah Risse, Eike Richter, & Thomas Weskott (in preparation)
Nachbarn machten grausigen Fund (NEUKÖLLN.)


3 Types of question:

... from the beginning: Ist die Polizeidirektion in der Hermannstraße? (Nein)

... from critical sentence: Fand der Nachbar den Mann? (Ja)

... from end: Wurde der Teppichhändler von seinen Kindern gefunden? (Nein)
Vocabulary over Age (N ~ 250)
Digit-Symbol Substitution over Age
Reading Time over Vocabulary

Kern, Weskott, & Kliegl (2008)
Comprehension over Age

Kern, Weskott, & Kliegl (2008)
Three Word Frequency Effects in Sentence Reading

Three Word Frequency Effects in Paragraph Reading

Kern, Weskott, & Kliegl (2008)
Three Word Frequency Effects in Paragraph Reading: Age Differences

Kern, Weskott, & Kliegl (2008)
Three Word Frequency Effects in Paragraph Reading: Digit-Symbol Substitution Differences

Worthäufigkeit

Kern, Weskott, & Kliegl (2008)
Three Word Frequency Effects in Paragraph Reading Vocabulary Differences

Kern, Weskott, & Kliegl (2008)
Lag Effects

Properties of word \( n-1 \) influence fixation duration on word \( n \)

Evidence from experiments manipulating target words:

- **preview benefit**
  e.g., Balota et al. (1985)

- **spillover**
  e.g., Rayner & Duffy (1986)

- **foveal load**
  e.g., Henderson & Ferreira (1990)
**Dynamical modulation of the perceptual span** (e.g., Henderson & Ferreira, 1990)

- the less frequent word $n-1$,
- the greater the foveal load during processing of word $n-1$,
- the more focused the perceptual span,
- the smaller the preview benefit for word $n$,
- the stronger the frequency effect on word $n$
For low-frequency word n-1, there is a stronger effect of the frequency of word n on the fixation on word n.
Word-level interactions for Fixation Duration

GLMM for model in Kliegl et al., 2006
Two 3-Factor Crosslevel Interactions for Fixation Durations

- Young adults
- Old adults
- Low-contrast young

Fixation duration [ms]

Low-freq word n
High-freq word n

Frequency of word n-1 [quantiles]

Graphs show fixation duration for young, old, and low-contrast young adults, with subplots for low-freq and high-freq words.
Effects of N-Frequency, N-1-Frequency, and Age on Fixation Duration in Reading Paragraphs

Paragraphs

Young adults | Middle adults | Older Adults

Frequency of word n-1 [quantile]

Fixation duration on word n [ms]

Rare word n

Frequent word n

Kern, Weskott, & Kliegl (2008)
III. Effects of Age on Scanpath regularity

IV. Reading strategy impacts on parafoveal-on-foveal effects in sentence reading

Parafoveal-on-Foveal (PoF) Effects

- Perceptual span: stronger parafoveal effects if word $n$ is short (Kennedy & Pynte 2005)
- Foveal load: more preprocessing of word $n+1$ if word $n$ is easy to process (Henderson & Ferreira 1990, 1993)
Jede Sprache der Welt besitzt eine Grammatik.
Selectivity effects in reading: 1. Age

- Reading is selective process: 10-30% of words are skipped

- Short, high frequent, predictable words are skipped more often (Rayner 1998, Drieghe et al. 2008): Function words are prime candidates for skipping

- In old age: more skipping, more regressive eye movements (Kliegl et al. 2004, Rayner et al. 2006, Laubrock et al. 2006)

In corpus analyses: Different reading behavior in terms of fixational selectivity (especially in old age) may lead to different composition of data base

May also give rise to critical parafoveal-on-foveal effects?
Martins gebrochener Zeh schwoll rasch an.

*Martin’s broken toe swelled quickly.*

<table>
<thead>
<tr>
<th>easy:</th>
<th>Was schwoll an?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What was swelling?</td>
</tr>
<tr>
<td>Fuß foot/ Ferse heel/ Zeh toe</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hard:</th>
<th>Was passierte mit Martins Zeh?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What happened to Martin’s toe?</td>
</tr>
<tr>
<td>wurde blau / wurde steif / wurde dick</td>
<td>became blue / became stiff / became thick</td>
</tr>
</tbody>
</table>
Questions and Measures

Reading for comprehension in 2 conditions and 2 age groups

- *Do reading intention and age affect the selectivity of fixated words?*

- *Is the selectivity of fixated words related to effects of distributed processing?*

**Variables of interest:** single fixation duration (SFD)

- selectivity in first-pass single fixation cases (SFC)

- effects of word frequency (n) on first-pass SFD

- effects of word frequency (n+1) on first-pass SFD
## Experiments

<table>
<thead>
<tr>
<th></th>
<th>easy young</th>
<th>easy old</th>
<th>hard young</th>
<th>hard old</th>
</tr>
</thead>
<tbody>
<tr>
<td>n subjects</td>
<td>24</td>
<td>32</td>
<td>30</td>
<td>23</td>
</tr>
<tr>
<td>Ø-age (years)</td>
<td>17.6</td>
<td>70.6</td>
<td>18.5</td>
<td>68.0</td>
</tr>
<tr>
<td>reading material</td>
<td>144 sent. (Potsdam Sentence Corpus)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>instruction</td>
<td>read for comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>freq. of questions</td>
<td>27%</td>
<td></td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>type of questions</td>
<td>easy</td>
<td></td>
<td>hard</td>
<td></td>
</tr>
</tbody>
</table>
Method

- **Eye Movement Recording:** EyeLink-II-System, 500 Hz

- Binocular calibration and recording

- Isolated sentences presented on the center line of a monitor, response via mouse klick

- **Analyses:** Effects of distributed processing on single fixation duration (SFD) with LMM (Pinheiro & Bates 2000)

- Separate models for old and young readers, focus on condition effects
## Results: summary statistics

<table>
<thead>
<tr>
<th>first pass</th>
<th>skipping</th>
<th>regression</th>
<th>SFD (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>easy young</td>
<td>.16</td>
<td>.07</td>
<td>231</td>
</tr>
<tr>
<td>hard young</td>
<td>.16</td>
<td>.18</td>
<td>242</td>
</tr>
<tr>
<td>easy old</td>
<td>.25</td>
<td>.14</td>
<td>224</td>
</tr>
<tr>
<td>hard old</td>
<td>.21</td>
<td>.25</td>
<td>245</td>
</tr>
</tbody>
</table>

- **age effect:** \(F(1,105)=25.2\) ***
- **cond. effect old:** \(F(1,53)=4.14\) *
- **age effect:** \(F(1,105)=7.2\) **
- **condition effect:** \(F(1,105)=27.8\) ***
- **condition effect:** \(F(1,105)=6.3\) *
## Results: selectivity effects (SFC)

<table>
<thead>
<tr>
<th>word n</th>
<th>frequency</th>
<th>length</th>
<th>predictab.</th>
<th>function word prop.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus-reference</td>
<td>2.3</td>
<td>5.4</td>
<td>-1.48</td>
<td>.37</td>
</tr>
<tr>
<td><strong>easy young</strong></td>
<td>2.24</td>
<td>4.5</td>
<td>-1.58</td>
<td>.32</td>
</tr>
<tr>
<td><strong>hard young</strong></td>
<td>2.26</td>
<td>4.5</td>
<td>-1.55</td>
<td>.33</td>
</tr>
<tr>
<td><strong>easy old</strong></td>
<td>2.03</td>
<td>4.9</td>
<td>-1.65</td>
<td>.27</td>
</tr>
<tr>
<td><strong>hard old</strong></td>
<td>2.23 ***</td>
<td>4.6 ***</td>
<td>-1.52 **</td>
<td>.33 ***</td>
</tr>
</tbody>
</table>
Summary results: selectivity (Single Fixation Cases)

- No selectivity effect between young samples.
- Demanding questions changed fixational behavior in old readers:

  Single Fixation Cases in easy old readers were proportionally more often:
  - lower frequent words
  - longer words
  - words of lower predictability

  hard old fixated more function words (FW) than easy old.
Word frequency effect: young

![Graph showing single fixation duration vs. frequency word n [log]](image)

- Linear: $t = -3.71$
- Quadratic: $t = -2.47$
- Cubic: $t = -8.31$

- Hard young
- Easy young
Parafoveal-on-foveal effect

\[ t = -5.56 \]

**Single fixation duration [ms]**

- **Easy young**
- **Hard young**

**Frequency word n+1 [log]**
Word frequency effect: old

- Difference: linear: $t = 4.05$, cubic: $t = -2.33$

- Single fixation duration [ms]

- Frequency word n [log]

- Hard old

- Easy old
Parafoveal-on-foveal effect

difference: $t = -3.84$

Single fixation duration [ms]

Frequency word n+1 [log]

hard old

easy old
Condition effects due to top-down strategy

- Fewer skippings, shorter sacc. amplitudes: due to reduced perceptual span in response to task demands
- Selectivity effect in old age: if fewer words are skipped, more FW are fixated (= short and highly frequent)

(Henderson & Ferreira 1990, Kennedy & Pynte 2005)

\[ \text{stronger next-word frequency effect on SFD} \]
Summary and Conclusion

• Reading strategy/ task demands impact in top-down fashion on eye movement behavior, reading identical material (Radach, Huestegge, Reilly 2008, Wotschack 2009)

• Selectivity effects in corpus analyses may give rise to PoF-effects
  — need to be considered in group comparisons

• Results confirm idea of distributed processing within the perceptual span

• Perceptual span modulated by reading intention and locally by processing demands