Measuring the Impact of Hyperlinks on Reading

Gemma Fitzsimmons, Mark Weal and Denis Drieghe
Reading for comprehension

- Most reading research focuses on reading for comprehension

- Theoretical and computational models of reading typically model the reading of single line text rather than passage text and focus to more on lexical processing than higher level issues (e.g. syntax, task effects)

- We are focusing on reading for comprehension with the additional task effects of reading on the Web
Why is it important to study reading on the Web?

- Users of the Web engage in a wide variety of different activities (Trend Data, 2012):
  - searching for information
  - reading the news/reading for comprehension
  - sending and receiving email
  - social networking

- Within all of these activities, the primary task that users engage in is reading text

- We want to explore reading for comprehension, but include the additional information of hyperlinks

- Does the extra high level information provided by hyperlinks effect reading behaviour?
Starting point – What are the main differences?

- The main differences between reading hypertext and offline reading are:
  
  - Hyperlinks are *coloured and salient* compared to the rest of the text. (Low-level influence)
  
  - There is *high level information attached to a hyperlink*. A hyperlink links one piece of information to another, perhaps on a separate page of the same Website, or a different Website.
How do we read hypertext?

- Very little research so far, mostly focused on Human Computer Interaction (HCI)
- Nielsen (1999) claims that “the mother of bad design conventions is the decision to make hypertext links blue”
- However, he also admits that hyperlinks being denoted in blue is now a convention and we should not move away from it
- Carr (2010) suggests that hyperlinks are a distraction in text and hinder comprehension
Saliency – Blue hyperlinks

- As far as we know no eye movement research on hyperlinks examining the effects of saliency on reading

- On the Web: Simola, Kuisma, Oorni, Uusitalo and Hyönä (2011) found that salient advertisements can distract attention and disrupt reading

- Off-line: White and Filik (2009) examined *italic* and **bold** words in passages of text. They found that bold text had shorter fixation durations. They suggest this reflects the improved visual discrimination of the text, perhaps making it easier to identify
High level processing – What is behind the hyperlink?

- High-level processes are those based on information not contained within the lexical representation of a word.

- Hyperlinks denote a connection to other content somewhere else on the Web.

- Closest thing to a prediction: E–Z Reader – Reichle et al (2009) suggest that higher-level processes intervene in eye movement control only when “something is wrong” and either send a signal to stop moving forward or a signal to execute a regression – exclusively impact later eye movement measures.
What are the main differences?

- The main differences between reading hypertext and offline reading are:
  - Hyperlinks are **coloured and salient** compared to the rest of the text
  - There is **high level information attached to a hyperlink**. A hyperlink links one piece of information to another, perhaps on a separate page of the same Website, or a different Website

- In order to explore both we ran two experiments on the **same participants**
Experiment 1

- *Does having a word in a different colour change reading behaviour?*

- 30 participants

- 5 conditions – target word is shown in either Black, Blue, Green, Red or Grey

The weather forecast said it would be sunny in England this weekend.
The weather forecast said it would be sunny in England this weekend.
The weather forecast said it would be sunny in England this weekend.
The weather forecast said it would be sunny in England this weekend.
The weather forecast said it would be sunny in England this weekend.
Experiment 1 Results

- Coloured words had reduced skipping compared to Black
Experiment 1 Results

- Grey (low contrast) words had longer fixation times compared to Black text (e.g. Drighe, 2008)

<table>
<thead>
<tr>
<th>Condition (Colour of Target Word)</th>
<th>Skipping Probability Percentage</th>
<th>First Fixation Duration</th>
<th>Single Fixation Duration</th>
<th>Gaze Duration</th>
<th>Go-Past Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>27 (19)</td>
<td>222 (36)</td>
<td>229 (54)</td>
<td>242 (47)</td>
<td>282 (106)</td>
</tr>
<tr>
<td>Blue</td>
<td>14 (15) **</td>
<td>219 (49)</td>
<td>225 (80)</td>
<td>244 (55)</td>
<td>280 (95)</td>
</tr>
<tr>
<td>Green</td>
<td>11 (14) **</td>
<td>234 (41)</td>
<td>243 (46)</td>
<td>260 (57)</td>
<td>339 (115)</td>
</tr>
<tr>
<td>Red</td>
<td>8 (10) ***</td>
<td>209 (36)</td>
<td>210 (44)</td>
<td>229 (47)</td>
<td>279 (59)</td>
</tr>
<tr>
<td>Grey</td>
<td>13 (13) **</td>
<td>244 (40) *</td>
<td>247 (51) *</td>
<td>279 (50) **</td>
<td>337 (96) **</td>
</tr>
</tbody>
</table>

Significantly different from Black = * $p<.05$, ** $p<.01$, *** $p<.001$
Discussion – *Does having a word a different colour change reading behaviour?*

- Experiment 1 – Decreased skipping of coloured words (compared to black)

- Yes, when a word is shown in different colour to the rest of the text it has decreased skipping, but it is not more difficult to process as fixation times remain constant (except grey)

- Good position to examine fixation times for hyperlinked words in Experiment 2, knowing that colour does not affect fixation times
Experiment 2

- *Does a hyperlinked word change reading behaviour?*

- *Additional question: Does it interact with lexical processing?*

- 32 participants – 4 conditions:

  Frequency

<table>
<thead>
<tr>
<th>Hyperlinked</th>
<th>Unlinked</th>
</tr>
</thead>
<tbody>
<tr>
<td>High High</td>
<td>Low Low</td>
</tr>
<tr>
<td>High Low</td>
<td>Low Unlinked</td>
</tr>
</tbody>
</table>
Train station

A train station is a railway facility where trains regularly stop to load and unload haulage or passengers. It consists of a platform next to the track and a station building providing services such as ticket sales and waiting rooms.

Railway stations usually have ticket booths or machines, or both, although on some lines tickets are sold on board the trains. Many stations include a convenience store providing drink facilities. Larger stations usually have fast-food or restaurant facilities. In some countries, stations may also have a bar or pub. Other station facilities may include: toilets, lost-and-found, departures and arrivals boards, waiting rooms and market facilities. There may also be taxi ranks and bus bays outside the station for easy transport links. A most basic station might only have platforms. Larger or manned stations tend to have a greater range of facilities but can also experience rodent problems.
## Experiment 2 Results

- **Significant effect of Word Frequency, Low Frequent words had longer fixation times**

- **No main effect of Hyperlinked/Unlinked**

<table>
<thead>
<tr>
<th></th>
<th>Skipping Probability Percentage</th>
<th>First Fixation</th>
<th>Single Fixation</th>
<th>Gaze Duration</th>
<th>Go-Past Times</th>
<th>Total Reading Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>High/Hyperlinked</td>
<td>43 (23)</td>
<td>216 (34)</td>
<td>214 (35)</td>
<td>227 (36)</td>
<td>298 (123)</td>
<td>261 (54)</td>
</tr>
<tr>
<td>High/Unlinked</td>
<td>46 (24)</td>
<td>215 (32)</td>
<td>219 (40)</td>
<td>228 (38)</td>
<td>291 (84)</td>
<td>266 (67)</td>
</tr>
<tr>
<td>Low/Hyperlinked</td>
<td>40 (24)</td>
<td>232 (45)</td>
<td>249 (47)</td>
<td>258 (47)</td>
<td>364 (112)</td>
<td>320 (68)</td>
</tr>
<tr>
<td>Low/Unlinked</td>
<td>43 (24)</td>
<td>231 (38)</td>
<td>238 (44)</td>
<td>251 (47)</td>
<td>306 (88)</td>
<td>297 (60)</td>
</tr>
</tbody>
</table>

| Frequency effect for Unlinked | -3 | 16 | 19 | 23 | 15 | 31 |
| Frequency effect for Hyperlinked | -3 | 16 | 35 | 31 | 66 | 59 |

| Difference between Unlinked and Hyperlinked | 0 | 0 | 16 | 8 | 51 | 28 |
The effect in total reading time is not due to re-reading after they have left the hyperlinked word, there was no difference in regressions in (but there was in regressions out)
Experiment 2 Results

- Linear mixed-effects model (lme) on untransformed data
  - Random effects: Participants and items
  - Fixed effects: Word frequency and Hyperlinked/Unlinked

- No decrease in skipping the hyperlinked (coloured) words that was seen in Experiment 1

- Significant interaction between Word Frequency and Hyperlinked/Unlinked in later measures indicating re-reading of the preceding text
Discussion – *Does a hyperlinked word change reading behaviour?*

- Experiment 2 – Longer times in later measures for Low frequent/Hyperlinked words
- Yes, hyperlinked words are processed differently to coloured words
- Extra high level information contained in a hyperlink
- Compatible with E–Z Reader (Reichle et al, 2009). Higher–level processes intervene in later measures only when the processing is difficult
- The high–level information of a hyperlink only causes re–reading when the word is low–frequent (more difficult to process)
Discussion

- The suggestion of a blue word representing a hyperlink in a Wikipedia environment is enough to show the same mechanism as a hyperlink would even without being able to click to hyperlinks.

- The low frequent hyperlinked words are perhaps the ones that we would be most likely to want to click.
Future Research

- Clicking

- Navigating through Webpages

- Other Webpages that are not Wikipedia – not all Webpages contain so many hyperlinks in the text where you can assume the destination is another similar Wikipedia page

- Task effects – reading for comprehension, searching for information
Thank you for listening!

Any questions?
## Appendices Experiment 1

### Eye movement measures for Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Skipping Probability Percentage</th>
<th>First Fixation</th>
<th>Single Fixation</th>
<th>Gaze Duration</th>
<th>Go-Past Times</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F (df)</td>
<td>F (df)</td>
<td>F (df)</td>
<td>F (df)</td>
<td>F (df)</td>
</tr>
<tr>
<td><strong>Main effect of colour</strong></td>
<td>7.06 (4,116)**</td>
<td>4.66 (4,116)**</td>
<td>2.30 (4,96)</td>
<td>6.10 (4,116)**</td>
<td>3.68 (4,116)**</td>
</tr>
<tr>
<td></td>
<td>t (df)</td>
<td>t (df)</td>
<td>t (df)</td>
<td>t (df)</td>
<td>t (df)</td>
</tr>
<tr>
<td>Black-Blue</td>
<td>3.19 (29)**</td>
<td>0.26 (29)</td>
<td>0.23 (27)</td>
<td>-0.23 (29)</td>
<td>0.09 (29)</td>
</tr>
<tr>
<td>Black-Green</td>
<td>3.42 (29)**</td>
<td>-1.65 (29)</td>
<td>-1.47 (27)</td>
<td>-1.73 (29)</td>
<td>-2.13 (29)*</td>
</tr>
<tr>
<td>Black-Red</td>
<td>5.18 (29)**</td>
<td>1.71 (29)</td>
<td>0.93 (29)</td>
<td>1.13 (29)</td>
<td>0.13 (29)</td>
</tr>
<tr>
<td>Black-Grey</td>
<td>2.80 (29)**</td>
<td>-2.61 (29)*</td>
<td>-2.56 (28)*</td>
<td>-3.43 (29)**</td>
<td>-2.79 (29)**</td>
</tr>
</tbody>
</table>
Appendices – Experiment 2

- $p$ values for the lme interaction – First fix = 0.45, Single fix = 0.08, Gaze = 0.13, Go–past = 0.02, Total time = 0.02

- Regressions results – Regressions IN – NON SIG

- Significant interaction in Regressions OUT. F(1,30)=4.448, p=0.043

- Means
  - Linked/High–17% – Unlinked/High–21%
  - Linked/Low–27% – Unlinked/Low–17%