Eye movement recording as a tool for studying language processing

Summary
Frenck-Mestre 2005
Rayner 2004, ms.
Clifton, Staub & Rayner (ms, review of 100 articles)
Measures (Rayner)

- **Reading (word based measures):**
  - First fixation duration
  - Single fixation duration
  - Gaze duration
  - Total fixation duration
  - Go-past time
  - Skipping
  - Regressions (in and out)

- **Reading (region-based):**
  - First pass reading time
  - Second pass reading time
  - Go-past reading time
  - Total reading time
Measures (Frenck-Mestre)

- First pass reading time
- Second pass reading time
- Total reading time (the sum of the two above)
  For any given ROI (region)
- **Gaze duration** – all fixations in the ROI from the first fixation until the eyes exit to the right or left
- Probability of skipping over an ROI
- Length of saccades (also informative)
- Pattern of regressions
- Probability of re-fixating a region
  A standard time frame – an average fixation is roughly 240 ms
Early vs. Late measures

- First pass time (and first fixation time)
- Second pass time (and total time to the extent that it reflects second pass time rather than first pass time)

Go-past and regressions-out measures - early / late:
- Occurrence of a regression reflects difficulty in integrating a word when it is fixated (an early effect).
- Go-past measure reflects this + also cost of overcoming difficulty which may occur late in processing.
General findings (Cliffton et al.)

- Not a clear picture at all (unlike lexical effects on word recognition)

- At first, it appeared that syntactic processing difficulty could be identified by quickly-appearing disruptions in the eyetracking record

- Later research, however, has not demonstrated such clear, immediate, and regular effects

- Also different results may come out of different measures
Survey of 100 eyetracking articles

• Measures of choice
  – FF (first fixation)
  – FP (first pass)
  – GP (go-past)
  – SP/TT (second pass or total time, whichever was reported)

• Warnings
  – Experiments differ in length of critical or disambiguating region (if it‘s very short, effects may only emerge on the following region)
  – Authors differ in how they divide sentences into regions – and this affects the emergence of effects …
So …

- "Answering this question requires a better understanding of the relation between comprehension difficulty and eye movements than we now have." (Clifton et al., ms)

- Some early research – syntactic and semantic anomaly slowed eye movements essentially immediately

- Some more recent – such anomalies may trigger regressive eye movements rather than affecting fixation durations

- Other research – effects of anomaly may in some cases appear later in the eye movement record
Is reading slower in the ambiguous region compared to unambiguous control condition?

• Models of sentence processing based on a time-consuming process of competition among alternative analyses predict such slowing –

• Very rare, however – sometimes even a speedup
• Mostly MC/RR (main clause – reduced relative) ambiguity

• Why?
  – Possibly undetected due to low power of experiment
  – Possibly some participants became aware of ambiguity and deliberately read it cautiously
Caution: do NOT analyze this

• Poor regions for analysis:
  – A) The beginning of a sentence
  – B) The end of a line of text prior to a line break
  – C) The end of a sentence

• Why?
  – (A) & (B) are skipped far more frequently than other sentence regions
  – (C) – end of sentence fixations – lo-o-ong!

• The “sentence wrap-up effect” (Just & Carpenter, 1980) -> readers spend more time at sentence final words not because of the word’s characteristics but because the entire sentence can be comprehended at this point
The effect of clause wrap-up on eye movements during reading

- Rayner, Kambe, & Duffy (2000)
  - Readers looked longer at a noun and the region after it when they were clause final
  - Readers regressed from a noun or the region after it more frequently when they were clause final
Caution again …

- **Eye movement patterns**
  - provide rich information
  - BUT only when compared to an equivalent, experimentally controlled alternative condition

- **Debate – distinguishing between kinds of factors:**
  - Low-level (psychophysical)
    - physical properties of words
    - the ocular-motor system
  - High-level (psycholinguistic)
    - Lexical frequency of words
    - Predictability of words
  - Low-level & High-level factors also interact – do NOT use short words as ROI, if possible
Caution 3 …

• “spillover effects” – linked to the processing of the preceding segment
Example: Altarriba et al. 1996

• Span-Eng bilinguals reading mixed or monolingual sentences
• Two factors – context vs. lexicon
  – Word frequency – high (1,2) vs. low (3,4) WF
  – Cloze probability/frequency (sentence constraint) – high (1,3) vs. low (2,4) CP
• Materials:
  – 1. She moved from the country to the city/ciudad to find a better job.
  – 2. We took a walk in the city/ciudad before we drove back home.
  – 3. On Halloween the children carved an orange pumpkin/calabaza for the front steps.
  – 4. the market had a new variety of pumpkin/calabaza in the fall.
• Measures (on target word)
  – Skipping probability
  – First fixation time
  – First pass gaze duration
• **Skipping probability**
  – more in high constraining context (on both mixed and monolingual sentences)
  – conclusion: context > lexicon

• **First fixations**
  – Monolingual sentences –
    • shorter for high constraint than low constraint
    • Independent of word frequency
  – Mixed sentences – interaction
    • Low frequency (Spanish) targets – high constraint facilitated processing
    • High frequency – low constraint facilitated processing (the opposite)

• **Gaze duration**
  – No significant results!