Towards an Annotation Standard for Eye Tracking Data

Michael Carl[1], Arnt Lykke Jakobse[1], and Oleg Spakov[2]
[1]*Copenhagen Business School, Computational Linguistics, ISV, Dalgas Have 15, DK-2000 Frederiksberg, Denmark, carl@iai.uni-sb.dk*
[2]*University of Tampere, Department of Computer and Information Sciences, Pinni B 1063, Kanslerinrinne 1, Tampere, Finland*

Eye-tracking is a growing field of research with several applications such as eye-movements in the recognition of natural environments, in human computer interaction, in speech and in reading. Eye-tracking has also been used as a means to understand the process of reading. This research focuses on the impact of single word properties, such as its duration, complexity (number of morphemes), familiarity of the word to the reader, its ambiguity, but also problems related to reading sequences of words in a context. Difficulties are usually measured by counting the duration of eye fixations on a particular word and the number of regressions to this word.

One of the aims within the Eye-to-IT project is to figure out what help a translator might need when reading a text with the goal to translate it into another language. It has been shown that reading - and thus Eye-fixation patterns - is very different in different reading purposes. Thus, [1] showed that reading a text of 200 words for understanding received 145 fixations, while there were almost 900 fixations when reading a similar text while typing a written translation of it. Besides the task of reading, fixation patterns also depend on the experience of the reader (or - for that purpose - on the experience of the translator), on the expected translation quality, on the difficulty and familiarity of the text to the reading subject, etc., etc.

Since many of the parameters causing particular fixation patterns are not well understood, we aim at creating a database of annotated texts, which can then be investigated in order to figure out the common patterns for different reading settings and to estimate the depending parameters.

The eye-tracking data and fixation patterns are obtained in an experimental setting, where a number of subjects are confronted with different reading tasks. An experiment supervisor would first introduce the subject to the task. The reading activities (and perhaps key-logging data, in case of written translation) would then be recorded by our software (Tobii 1750 eyetracker and Translog/GWM). In a third step the supervisor would go through the data together with the subject to assemble additional information.

This last step resembles a 'think aloud' protocol, with the difference that subjects see their own gaze activities on the screen and comment on previous actions. A subject would then explain where and what kinds of problem were faced during the reading task and what strategies he/she followed to solve those problems. The revision session also serves to readjust eye-mapping errors which occur when mapping the eye position data on a word. The tool, thus consists of the four following steps:

Collect **eye-position data,** i.e. the raw output data from eye-tracking devices. The TOBII 1750, for instance, produces a data sample every 20ms (at 50 Hz) consisting of the pixel positions (X,Y) for the left and the right eye and their pupil dilation

Eye-fixations are computed from the eye-position data and mapped on the characters and/or words looked at on the screen. This so-called **eye-mapping data** maps the eye-fixation data on (sequences of) symbols similar to the d*rift correction algorithms* [2]. Due to the inaccuracy of the technical devices and the visual system, this process is, however, not always correct. This is particularly so, if the reading task involves searching the screen where the gaze moves back and forth across different regions on the screen. A manual re-adjustment is therefore in many cases necessary.

During **mapping re-adjustment** the experiment supervisor and the subject go through the text that was read, together, and revise the eye-mapping data. Mapping re-adjustment is made possible with a tool in which the eye-position data is shown together with the eye-mapping data during a revision session. An example of a possible screen shot is shown in Figure 1. The dots show the eye-position data. Their size indicates the duration of the fixation, while the highlighted characters show the eye-mapping data. It is possible to review the reading session in real time, to slow it down, and to manually adjust erroneous eye-mapping positions. In addition, comments of the subject may be added, which reflect his/her mental state at a given moment.

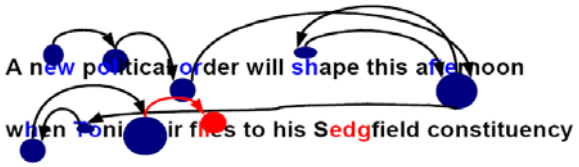

***Figure 1**: Sketch of the Annotation Tool.*

These **speak aloud annotations** will be stored with the mapping data in a separate field. In addition to 'free annotations', the subject will also be able to choose from a set of previously anticipated categories of mental states which best describe the subject's mental state at the given moment. The aim of this annotation is to enable statistical investigation of the annotated material in order to find typical patterns which lead to or entail the mental state.

The described annotation tool aims at 1. elaborating a standard for eye-movement annotation in the reading task, 2. making it possible to collect a substantial amount of comparable eye-movement data in order to 3. statistically investigate similarities and differences in eye-movement patterns across different reader, reading tasks and reading conditions and 4. to share our data, results and tools with other researcher.

References

1. Jakobsen, A. L. and K. T. H. Jensen (2007). Coordination of comprehension and text production in witten and oral translation tasks. Poster presented at the 13th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP) 2007.
2. Hyrskykari, A. (2006) Utilizing eye movements: Overcoming inaccuracy while tracking the focus of attention during reading. *Computers in Human Behavior* **22**(4), 657-671.