

Functions and Objectives of a Query Language for User-Activity Data

Michael Carl

Department of International Language Studies and Computational Linguistics ,
2000 Frederiksberg, Denmark

One of the aims of the Eye-to-IT project (FP6 IST 517590) is to integrate keyboard logging and eye-tracking data to study and anticipate the behaviour of human translators. This so-called User-Activity Data (UAD) would allow to empirically ground cognitive models of translation and to validate hypotheses of human processing concepts in the data.

Two programs have evolved in the Eye-to-IT project: TRANSLOG is designed to register and replay keyboard logging data, while GWM is a tool to record and replay eye-movement data. The talk reports on an attempt to synchronise and integrate the representations of both software components so that sequences of keyboard activities and eye-movement data can be retrieved and correlated with textual segments they are associated with. The talk will raise question concerning the basic functions and objectives of a UAD query language.

UAD can be studied from three different angles: 1. Patterns of fixations, 2. Keyboard activities and 3. Linking both modalities.

Patterns of fixations are related to the process of text understanding. It will therefore be interesting to elaborate a taxonomy of fixation patterns and correlate phenomena of text understanding with it. Keyboard activities, on the other hand, give insight into the text production process. A taxonomy of keyboard patterns will thus provide a framework to classify text production mechanisms. Finally, linking related source text fixation patterns with patterns of keyboard activities unveils peculiarities of the actual translation process.

Orthogonal to these three angles of study, UAD could be queried with a view from the product data so as to ask what kinds of process data are related to particular set of textual data. For instance, the query language would allow to describe sequences of words in the source or/and in the target text with the aim to retrieve, compare and quantify associated patterns of fixations or/and keyboard actions.

On the one hand, UAD could be queried with a view from the process data in order to retrieve and compare associated textual data. The query language would, thus, allow to describe patterns of fixations or keyboard activities and to retrieve associated sequences of text. It thus becomes possible to study the interplay of fixation and keyboard pattern and investigate their causal relations in time. The talk will give a few examples of such retrieved patterns.