

Olli Philippe Lautenbacher
Centre for Translation
University of Turku, Finland
(oplauten@utu.fi)

Eye-to-IT Conference

Multimodality and Saliency– When Eye-Tracking Meets Subtitling

Film is multimodal by nature, combining multiple semiotic codes through visual image and sound. Furthermore, in subtitled film, two different kinds of visual information co-occur with the original soundtrack, i.e., the pictorial filmic elements and the subtitles' textual elements. Hence, from a translational viewpoint, the question of relative saliency of these co-occurring elements, i.e., what is primarily perceived by the spectator, might be more important than expected. A number of questions arose in basic eye-tracking tests conducted at the French Translation Department, at Turku University. These tests were made on still pictures combining text and image in different manners. The size of the two subject groups was only of 9 and 10 translation students, which of course limits the validity of the results. Nevertheless, the tests led to some hypotheses on future eye-tracking tests on subtitle-reading in film documents we are planning to conduct. I will present some of these hypotheses here and discuss them further at the conference.

In linear text, many studies underline our propensity to begin reading at the upper left corner of a text. They also show how readers jump from a given gaze location to another, following certain tendencies, not having systematic fixations on every word for example. On the other hand, visual image also carries some natural tendencies concerning the relative saliency of its components, such as the fact that human faces and their gaze in pictures are very powerful gaze-catchers to the viewer. These phenomena in reading and picture viewing were easily observed in the corpus. But what happens when these natural saliencies meet?

First, perceptual saliencies can be in contradiction with one another. But are there any primary areas of interest in the filmic picture that could systematically alter the natural reading of the subtitles (or vice versa)? The eye-tracking tests suggested a human propensity (within student subjects) to read what is on screen to be read (to a certain extent) as well as fixating faces and gazes.

Secondly, these saliencies can also be cohesive and work together. The visual-verbal cohesion system suggested by Baumgarten (2008) shows how different kinds of parallels occur between (spoken) text and exophoric references in the filmic picture. Correspondingly, another question is whether these "meeting points" between specific pictorial objects and textual elements, simultaneously on screen in subtitled films, would result in a stronger saliency of those semiotically connected elements, thus creating a specific dynamic move in the spectator's gaze? And does this double visual saliency shadow other potential areas of interest?

Thirdly, a methodological aspect also observed in the tests and worth underlining here is the primary importance of the distance of observation and screen size, especially when the eye-tracking analysis concerns relative saliency of visual elements. These factors differ radically between computer, television and cinema screens. The tests showed clearly that the visual areas covered by parafoveal vision in each situation need to be thoroughly taken into account in eye-tracking studies in general and on subtitling in particular.

These are central questions to be taken into account in practice, for instance in the physical (both spatial and temporal) positioning of subtitle on screen.