Doing multimodal corpus linguistics: investigating changes in Helsinki tourist brochures between 1967 and 2008

Tuomo Hiippala
Department of Modern Languages, University of Helsinki, Finland

This poster presents the results of a doctoral dissertation investigating the multimodal structure of tourist brochures published between 1967 and 2008 by the city of Helsinki, Finland. The aim of the research project was to develop a framework for modelling the structure of printed multimodal artefacts.

Drawing on an empirically-motivated framework, the research project annotated and compiled a multimodal corpus of 58 double-pages from a total of 30 different brochure series published during the studied period of time. The brochures were accounted for their content, layout, rhetorical, and navigation structure using a cross-referenced, XML-based annotation scheme.

To assist in the search for patterns in the data, the dissertation developed novel visualisation methods, which combined data from multiple analytical layers for a visual representation of the brochures’ multimodal discourse structure. This allowed the project to build bridges between the theoretical framework and the observed multimodal structures, achieving research results that would not have been possible without a corpus-based approach and the developed tools.

Measuring the proportions between language and images in the tourist brochures showed that the tourist brochures have been visual throughout the studied period of time, as the average amount of layout space taken up by images was 50,50%. The discourse structure of the brochures, however, had undergone a significant change after the year 1985, which also marked the introduction of desktop publishing software. This change, which involved a shift from linear-interrupted written language towards a fragmented non-linear design, was identified by analysing the structural patterns found in the corpus. This difference proved also statistically significant.

The dissertation underlined the benefits of doing empirical research on multimodality and using visualisations in analysis, while also identifying a major challenge. Without computer-assisted annotation, the production of sufficiently large corpora will remain time- and labour-intensive, thus limiting both theoretical and practical development.

Keywords: corpus, multimodality, visualisation.