The e-cosmos project - Integrating electronic tools to derive cognitive, linguistic and behavioral patterns from complex multimodal databases in the humanities

Stella Neumann and Paula Niemietz, RWTH Aachen University

Complex multimodal bodies of data such as text production behavior data and combined speech and communicative action data represent a new frontier in humanities research demanding novel approaches to the mathematical modeling, mining and quantification of these data. The data involved comprises synchronous streams of eye-tracking and keystroke logging data as well as data obtained from recording speech production and motion capture of gestures and other body movements. Because of its complexity, this type of multimodal data is usually only investigated in small-scale exploratory studies.

The interdisciplinary project team consisting of data mining experts, experts in information management in mechanical engineering and linguists will process these combinations of data to detect cross-modal clusters as well as hidden patterns in multimodal and time-displaced data streams. Drawing on an adaptive architecture developed in the context of production technology, an integrated set of tools will be compiled with the help of user-centered design. This platform will be used by the linguists to undertake quantitative, yet fine-grained linguistic investigations into the cross-modal structures.

This presentation will give an overview of the two linguistic case studies from translation-related text production and the combined production of speech, gesture and facial expressions and introduce the computational issues in processing the data for the purpose of quantitative analyses.