

The notion of salience in the discourse analysis models how the activation of referred objects evolves in the flow of text. The salience algorithm was already defined and tested briefly in an earlier research, we present a reproduction of its results in a larger scale using data from the Prague Dependency Treebank 3.0. The results are then collected into an accessible shape and analyzed both in their visual and quantitative form in the context of the two main resources of the salience – coreference relations and topic-focus articulation. Furthermore, a possibility of modeling the salience degree by a machine learning algorithm (decision trees and random forest) is examined. Finally, attempts are made with using the salience information in the machine learning NLP task of document clustering visualization.