Computational Investigation of Linguistic Markers in Discourse of Political Adversaries via Interpretation of Recurrent Neural Network

Kirillov B.A.*, Petrovskaya A.V.**, Sergeeva A.S.***

* Skolkovo Institute of Science and Technology, Moscow, Russia; ** National Research University Higher School of Economics, Moscow, Russia; *** ITMO University, Saint Petersburg, Russia

8k1r1ll0v@gmail.com; alexvalpetrovskaya@gmail.com

Cognitive-discoursive paradigm puts forward discourse of a community as the major unit under analysis. A community discourse can be analyzed through texts written by participants of the community since it is expressed in those texts in various ways. For large communities, sheer amount of texts generated limits the ability of human researcher to comprehend unique features of a discourse. But modern Machine Learning algorithms are able to process large amount of texts thus aiding the human researcher in investigation. We offer a modern word-level Recurrent Neural Network-based approach for unsupervised detection of discourse-specific linguistic markers: specific phrases, word sequences, metaphorical models and separate words. The first phase of our research is collecting a large corpus of three text types: writings of Russian pro-government and opposition activists and neutral texts without political coloring. The neutral corpus serves as control group. The second phase is training the model to classify the texts until it reaches near-human performance (measured by comparing its output to labels made by human participants). And the final phase involves interpretation of trained model to highlight the exact kinds of words that model links to a class. We suggest our approach for revealing implicit intents and linguistic markers and claim that it is universal for any kind of discourse.

Keywords: neural networks, political discourse, linguistic markers.