

MULTIMODALITY IN TOPIC MODELING

Nadiia Doronkina

Senior Lecturer,

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Modality is referred to as a phenomenon connected with perception of information. In particular, this term is used to signify the sensory channels of perception. In cognitive linguistics it is defined as one of information codes included in the concept of hybrid text. In psycholinguistics modality means membership in a certain sensor system expressed in characteristics of senses, stimuli, information and sensory organs. So, there are the following types of modalities: smelling, touching, tasteful, sighting and hearing. Therefore, multimodality involves such modi as writing, speaking, images to create meanings. The function of multimodality is to describe the interaction rules of verbal and nonverbal signs within a communicative act (Krejdlin, 2014, p. 101; Sorokina, 2017, p. 168).

Any discourse is multimodal because the process of communication implies interaction of different semiotic systems, multimodality serving as the way to organize the system for the purpose of senses production. Multimodal approach to communication research provides a realistic picture of natural language communication taking account of all communication channels and their internal organization. In corpus linguistics a number of multimodal corpora including text, audio and visual materials are built (Kibrik, 2010, pp. 147-148).

In terms of multimodal approach discourse is analyzed at three levels:

- microlevel means the division of documents into different codes and subcodes resulting in detecting the resources involved in communicative process of each semiotic system;
- mesolevel considers interaction between microlevel elements, supplementary meaning potential modifying the sense of communication;
- microlevel covers both each code potential and their mutual activation (Sorokina, 2017, p. 169).

The applications of multimodal discourse analysis include topic modeling. In particular, that approach is widely used in statistical analysis to solve such tasks as recognition of the trends in archives of scientific published works, patent databases; information search; classification and categorization of documents; topic text segmentation; image and video stream analysis etc.

Multimodal topic model describes documents that contain metadata along with main text. On the one hand, metadata redounds to specifying the topics of documents. On the other hand, topic models are used in order to detect semantics of metadata or foresee the metadata missed. Each type of metadata creates its own modality and vocabulary, for example such text modalities as word combinations, tags, references. Besides, the modality of letter n-grams is used to analyze short texts containing literal errors with the purpose to improve information search quality. The example of non-textual modalities is graphical elements of images. A multimodel topic model extends semantics of topics to its elements of other modalities including nontextual ones (Bol'shakova et al., 2017, p. 218).

Topic modeling deals with big collections of documents (e. g. scientific articles in a journal, news stream, books etc.) in order to identify hidden semantic structures. A common topic model specifies topic distribution in documents and documents in topics.

The process of task solving is divided into the following stages:

- 1) transformation the input data into the appropriate format;
- 2) creating the model;
- 3) optimization the model;
- 4) evaluation;
- 5) actual solution.

Any text in topic models is compressed and represented in the form of a vector where each coordinate corresponds to a certain topic to which the list of key words and phrases describing semantics of that topic is attached. In processing the document is considered as a “bag of words” (i.e. the number of words occurrence is calculated disregarding their order) (Sivic, 2009).

The idea of topic modeling is based on the hypothesis that any document contains a unique blend of topics with different degrees of importance. Creating the model it is necessary to specify its desirable characteristics:

- the needed number of topics, the rest of them being considered to be unimportant;
- the features of modalities (their names and coefficients of importance);
- optimization criteria, so called regularizers.

Additive Regularization for Topic Modeling (ARTM) makes it possible to set several criteria of regularization at the same time. BigARTM is a library used as a tool to infer topic models. The regularizers realized there perform such functions as smoothing, rarefication, decorrelation and selection. Smoothing helps to separate background topics and words. Rarefication results in detection of the topics in a document and lexical nuclei of topics in a vocabulary. Decorrelation is used in order to differentiate lexical nuclei of topics. Selection of topics causes insufficient ones to be ignored. The output of model creation consists of the three parts: 1) the resulting set of topics, each of them is a group of words; 2) the matrix Φ describing topic word distribution; 3) the matrix Θ describing document topic distribution. In the process of model concordance it is changed with respect to regularizers which are applied according to the predetermined strategy. In optimal model the matrices Φ and Θ are rarefied (i.e. most elements are zero). It means that topics are interpreted, i.e. they don't contain grammatical and frequently used words.

Topic model evaluation is performed with the help of internal and external criteria. The former assesses the model based on Φ and Θ . The latter analyzes the model's ability to do its task (Blei, 2012).

References

Blei, D. M. (2012). Probabilistic topic models. *Communication of the ACM*. 55(4), 77-84.

Bol'shakova, Ye. I., Voroncov, K. V., Yefremova, N. E., Klyshinskij, E. S., Lukashevich, N. V., & Sapin, A. S. (2017). *Avtomaticheskaya obrabotka tekstov na yestestvennom yazyke i analiz dannyh: Uchebnoye posobie [Automatic processing of*

texts in natural language and data analysis: study guide]. Moscow: NIU VSHE. [in Russian].

Kibrik, A. A. (2010). Mul'timodal'naya lingvistika [Multimodal linguistics]. *Kognitivnye issledovaniya [Cognitive research]*, 4, 134-152. [in Russian].

Krejdlin, G. Ye. (2014). Semioticheskaya kontseptualizaciya tela i problema mu'timodal'nosti. [Semiotic conceptualization of body and problem of multimodality]. *Ekologiya yazyka s kommunikativnaya praktika [Language ecology and communication practice]*, 2, 100-120. [in Russian].

Sivic, J. (2009). Efficient visual search of videos cast at text retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 31(4), 591-605.

Sorokina, Yu. V. (2017). Ponyatiye modal'nosti i voprosy yazyka mu'timodal'nogo lekcionnogo diskursa [The concept of modality and questions of multimodal lecture language]. *Filologicheskiye nauki: Voprosy teorii i praktiki [Philological sciences: Theory and practice]*, 10(76), 168-170. [in Russian].