

# Targeting in Media Communication – Cognitive Models and Computational Applications

Peter Uhrig<sup>1</sup>, Irina Pavlova<sup>2</sup>, Scott Hale<sup>3</sup> & Anna Wilson<sup>2</sup>

<sup>1</sup>Technische Universität Dresden, peter.uhrig@tu-dresden.de <sup>2</sup>Oxford Internet Institute, University of Oxford, <sup>3</sup>Oxford School of Global and Area Studies, University of Oxford

**Keywords:** targeting (cognitive process), media communication, multimodality, empirical linguistics, automation

One major issue for multimodal research in cognitive linguistics is how to be certain that viewers of media broadcasts construct meaning the way that analysts claim they're likely to do. We address this question by analysing the relationship between RT videos and audiences' comments to them. In a first step we thus need to establish the points of reference in the corresponding video (or beyond). We identify these drawing upon Talmy's (2018) cognitive mechanism of targeting. This is time-consuming to do manually, and thus we need to engage with another big question in our field at the moment, viz. how we can scale up our analyses in order to make them statistically valid and reliable.

Our analysis is based on 20 videos from RT's show SophieCo Visionaries and the top-level comments (i.e. we excluded comments replying to other comments) collected before RT was banned from YouTube in early 2022. The following examples were found under a video discussing 5G technology:

- (1) Cant wait to perform brain surgery from home
- (2) Extremely IGNORANT. 5g is 60 hz. 60 hz is the same as radiation therapy for cancer patients. People are fuckin stupid. No one needs more than 4g lte.
- (3) May Jesus have mercy on your souls

Example (1) can be linked to a segment in which the show's guest explains how low-latency networks enable applications such as remote surgery. The situation is trickier with example (2), which does not reference anything specific in the video but rather the overall topic of the show. The type of comment exemplified by example (3) cannot be reliably related to anything in the video even by humans.

We discuss ways in which this human interpretation can be implemented, mirrored, or at the very least supported by software. To this end, the prototype of a system is evaluated that runs speech recognition with Whisper (Radford et al. 2022) on the videos, splits the results into sentences, and applies coreference resolution (Clark and Manning 2016). The results are then processed with a baseline system based on BM25 text similarity (Robertson & Zaragoza 2009) and a more advanced system that uses sentence embeddings fine-tuned on question answering systems (Reimers and Gurevych 2019) to compare the sentences from the video to the comments.

In the process of evaluating the results of our computational analysis we assess the compatibility of concepts and methods coming from cognitive linguistics and computer science. We don't just consider the cognitive mechanism of targeting, but also conceptual (viewpoint) blending (Fauconnier and Turner 2002) and stance-taking (Du Bois 2007). In doing so, we shed further light on the cognitive modelling of targeting as based on our data.

Finally, we will suggest for future research an extended system that goes beyond the text and also relies on automatic object recognition in the video to find further potential targets for the comments.

**Acknowledgement:** This research was generously funded by the Deutsche Forschungsgemeinschaft (Germany) and by the Arts and Humanities Research Council (UK).

## References

- Clark, Kevin and Christopher D. Manning. 2016. Deep Reinforcement Learning for Mention-Ranking Coreference Models. In *Proceedings of EMNLP*.
- Du Bois, John W. 2007. The stance triangle. In: Robert Englebretson (ed.) *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139-182. Amsterdam: John Benjamins.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. New York: Basic Books.
- Reimers, Nils and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics.
- Radford, Alex, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey and Ilya Sutskever. 2022. Robust Speech Recognition via Large-Scale Weak Supervision.  
<https://arxiv.org/abs/2212.04356>
- Robertson, Stephen and Hugo Zaragoza. 2009. The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval* 3(4).
- Talmy, Leonard. 2018. *The Targeting System of Language*. Cambridge, Massachusetts: MIT Press.