
Aspectual Class Classification as a Linguistic Approach to Event Detection in Natural Language Texts

Theresa Krumbiegel

*Fraunhofer Institut für Kommunikation, Informationsverarbeitung und
Ergonomie (FKIE)*

theresa.krumbiegel@fkie.fraunhofer.de

Recent event research in computational linguistics skews heavily towards content classification tasks like event type classification and event argument extraction. In these contexts, the layperson definition of the term *event* is often used, in the sense of *newspaper-worthy happening*. The linguistic definition of events is usually disregarded. According to linguistic theory, features like telicity and dynamicity can help to determine the aspectual class of a verb, which in turn indicates the presence or absence of an event in a sentence. Even a simple sentence like *John painted a picture in an hour* (Dowty, 1979, p. 56) can be classified as an event according to Vendler (1967), who differentiates between four aspectual classes: states, activities, achievements and accomplishments. Human annotators can identify and label these classes with the help of a number of criteria as summarized by Dowty (1979). The question remains whether transformer models such as BERT (Devlin et al., 2019) are able to: (1) classify sentences according to these classes while (2) basing their classification decisions on relevant features in the sentences. Some experiments regarding the automatic classification of aspectual verb classes have already been conducted (e.g., Metheniti et al., 2022). However, the data used for these experiments is often engineered to better suit the task. Sentences follow a simplified syntactic structure, thereby artificially avoiding the problem of aspectual type coercion (Moens & Steedman, 1988) as well as possible (semantic) bias in the data base. The current study aims to assess how well transformer models can perform this classification on a more diverse dataset. A corpus of online newspaper articles will be manually annotated using Dowty's criteria as a guide. This dataset will then be used to conduct a series of proof-of-concept classification tasks: two binary classification tasks to detect telicity and dynamicity respectively, and a multiclass classification task focusing on Vendler's aspectual classes.

References: • Dowty, D. (1979). *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel. • Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY. • Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the NAACL: Human Language Technologies, Volume 1*, pp. 4171–4186, Minneapolis, MN. ACL. • Metheniti, E. et al. (2022). About Time: Do Transformers Learn Temporal Verbal Aspect? In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pp. 88–101, Dublin, Ireland. ACL. • Moens, M. & Steedman, M. (1988). Temporal Ontology and Temporal Reference. *Computational Linguistics* 14(2), 15–28.