AI-DRIVEN AUTOMATIC DETECTION OF METAPHORS AND IDIOMS IN ENGLISH

Isomova Feruza Nizom qizi
Master's Alumna of Webster University in Tashkent

Abstract: *This study explores AI-driven approaches for automatically detecting metaphors and idioms in English by integrating linguistic theories with modern NLP technologies. It highlights how frameworks such as MIP, SPV, and Cognitive Metaphor Theory support the identification of figurative expressions in context. The paper analyzes the evolution from rule-based methods to deep learning and Large Language Models that achieve high accuracy in figurative-language classification. Remaining challenges and future directions, including contextual understanding and embodied cognition, are also discussed.*

Keywords: *Metaphor detection; Idiom identification; Figurative language; NLP; Deep learning; Large Language Models; Cognitive linguistics.*

The automatic detection of metaphors and idioms in English has become an increasingly significant area of research within Natural Language Processing (NLP), particularly in the context of rapidly advancing deep learning methodologies and Large Language Models (LLMs). As figurative language represents a central aspect of human communication, the development of computational tools capable of distinguishing figurative from literal expressions is vital for enhancing machine understanding, improving human–computer interaction, and supporting a range of linguistic and cognitive studies. This field is inherently interdisciplinary, combining insights from linguistic theory, cognitive science, and machine learning to address the challenges posed by the contextual, cultural, and conceptual complexity of figurative expressions. From a linguistic perspective, metaphorical and idiomatic expressions are notoriously difficult to process because they exhibit non-compositionality and context sensitivity: their meanings cannot be derived solely from the literal meanings of their constituent words but rather emerge from conventionalized or contextually novel interpretations. One influential analytic framework is the Metaphor Identification Procedure (MIP), which assesses metaphoricity by identifying a semantic gap between a word's contextual meaning and its basic, historically grounded meaning. If the contextual meaning contrasts with but remains relatable to the basic meaning through similarity or analogy, the word is classified as metaphorical. Another important linguistic principle is Selectional Preference Violation (SPV), which occurs when a word disrupts the expected semantic constraints of its surrounding lexical environment—for instance, the verb twist typically selects for physical objects, making the phrase "twist my words" a likely candidate for figurative interpretation. These linguistic diagnostics are grounded in broader theoretical frameworks, particularly Cognitive Metaphor Theory, which views metaphors as fundamental cognitive mechanisms that structure how humans conceptualize abstract domains through more concrete, embodied experiences. This theory has informed computational models by encouraging the integration of semantic variables such as concreteness, imageability, emotional valence, and conceptual mappings.

Technologically, approaches to automatic metaphor and idiom detection have evolved substantially over the past two decades. Early systems were predominantly rule-based and relied on manually encoded knowledge, lexical databases such as WordNet, and shallow statistical methods. Although groundbreaking at the time, these models were limited by their lack of scalability, dependence on expert-crafted resources, and difficulty generalizing across domains. The shift toward machine learning and deep learning introduced more robust mechanisms for pattern recognition, enabling systems to automatically learn discriminative features from large corpora. Neural architectures such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks became foundational for this task, often supplemented by linguistic features including part-of-speech tags, syntactic dependencies, and lexical concreteness norms. More advanced hybrid architectures, such as CNN+SVM or Bi-LSTM+CNN models, emerged to combine local and global contextual information, yielding notable improvements in classification performance.

The advent of Transformer-based architectures, particularly BERT, RoBERTa, GPT-style models, and other LLMs, has profoundly reshaped the landscape of figurative language processing. These models acquire deep contextual representations by pre-training on massive corpora, enabling them to capture nuanced semantic and syntactic patterns. When fine-tuned on annotated metaphor and idiom datasets, LLMs often achieve state-of-the-art results, at times approaching or surpassing human-level performance on controlled benchmark tasks. Techniques such as prompt engineering, few-shot learning, and Retrieval-Augmented Generation (RAG) further enhance the models’ ability to interpret figurative expressions by incorporating external knowledge and enabling dynamic reasoning.

Despite these advancements, several challenges persist. Figurative language is highly dependent on context, discourse, and socio-cultural background, and current AI systems frequently struggle to model these broader interpretive dimensions. Moreover, humans draw on common-sense reasoning and embodied cognition—including sensory experience, physical interaction with the world, and analogical thinking—to interpret metaphors, capacities that AI lacks by virtue of its disembodied nature. Another critical limitation concerns data availability: large-scale metaphor and idiom annotation requires extensive human effort, and annotators themselves often disagree, resulting in inconsistencies that hinder the training of robust models.

Looking ahead, future research is expected to explore models that integrate neuro-cognitive evidence, such as brain imaging data, alongside theories of embodied cognition, with the aim of more closely approximating human interpretive processes. Multimodal learning, grounding language models in perceptual data, and developing richer cross-cultural corpora are also promising directions. Ultimately, advances in metaphor and idiom detection will depend not only on algorithmic innovation but also on deeper interdisciplinary collaboration to bridge the gap between computational efficiency and human cognitive complexity.

REFERENCES:

1. Automatic Metaphor Detection using Large-Scale Lexical Resources and Conventional Metaphor Extraction Yorick Wilks, Adam Dalton, James Allen, Lucian Galescu (2013) https://aclanthology.org/W13-0905/

2. Cross-Lingual Metaphor Detection Using Common Semantic Features Yulia Tsvetkov, Elena Mukomel, Anatole Gershman (2013) https://aclanthology.org/W13-0906/

3. MIss RoBERTa WiLDe: Metaphor Identification Using Masked Language Model with Wiktionary Lexical Definitions

Applied Sciences (2022) https://doi.org/10.3390/app12042081

4. A Systematic Review of Figurative Language Detection: Methods, Challenges, and Multilingual Perspectives Natural Language Processing Journal (2025) https://www.sciencedirect.com/science/article/pii/S2949719125000688

5. Methods of Annotating and Identifying Metaphors in the Field of Natural Language Processing Martina Ptiček & Jasminka Dobša (2023) https://doi.org/10.3390/fi15060201

6. Metaphor Detection with Context Enhancement and Curriculum Learning Kaidi Jia & Rongsheng Li (NAACL 2024) https://aclanthology.org/2024.naacl-long.149/