Crossref doi | Google Scholar | WorldCat | MENDELEY

OSCAR PUBLISHING SERVICES

Research Article

INVESTIGATING THE SEMANTIC FEATURES OF SYNONYMS IN PARALLEL CORPORA

Pulatova Niso Shamuratovna

A teacher of Termiz State University, Uzbekistan

ABSTRACT

Synonyms, or words with similar meanings, are critical components in natural language processing (NLP) activities. Understanding their semantic properties, especially in parallel corpora, offers insights into translation, sentiment analysis, and information retrieval systems. This article explores the semantic complexities of synonyms in parallel corpora, emphasizing their importance and ramifications.

KEYWORDS

Parallel corpus, corpus, synonyms, synonymy, synonym features, semantic features.

INTRODUCTION

Synonyms are abundant in any language, posing both a challenge and an opportunity for NLP researchers and developers alike. While synonyms may appear to be interchangeable at first look, tiny changes in meaning, connotation, register, and usage can have a substantial impact on the accuracy and effectiveness of language processing systems. For example, in machine translation, finding the best appropriate synonym might be critical to creating contextually accurate translations. Similarly, in sentiment analysis, recognizing the specific emotional connotations and minor differences in sentiment represented through synonyms can result in more fine-grained and accurate sentiment classification.

Parallel corpora, which are composed of aligned texts in multiple languages, are a valuable source of linguistic data for studying the semantic properties of synonyms. By analyzing synonymous word pairs.

METHODS

Synonyms' semantic qualities include lexical, syntactic, and contextual dimensions. In logical semantics (also referred to as analytical semantics), semanticists depend on synonymy in order to prove the truth of a statement.   In parallel corpora, these characteristics show as translation equivalence, divergence, and convergence.

RESULTS

Among the approaches used in the recognition of synonyms, substitution has been one of the most persistent criteria (Palmer, 1981; Lyons, 1981; Cruse, 1986; Hoey, 1991; Sinclair, 1991; Stubbs, 2001). Traditionally, two words are considered synonymous in a sentence or linguistic context if the substitution of one for the other does not alter the truth value of the sentence. This explanation has however been shown to be not only ambiguous but impractical in determining whether candidate

words are synonyms or not Likewise componential analysis has also proved

ineffective in defining synonymy and discriminating between synonymy and co-hyponymy. Despite the deficiency of substitution and componential analysis in differentiating synonymy, no other approaches were proposed until the development of corpus linguistics. Since then, although a number of corpus investigations have been conducted into synonymy by looking at their collocations and semantic prosodies, there have been few holistic or systematic studies of synonyms.

DISCUSSION

1. Parallel corpora are a valuable resource for studying translation equivalents of synonyms between languages. By aligning corresponding portions, researchers can examine how synonyms are translated and detect small semantic changes. For example, while "happy" and "joyful" are frequently translated identically, contextual nuances may determine their use in different situations.

1. Translation Equivalence:

• "House" and "home" may be translated interchangeably in some contexts, but subtle differences exist. For instance, "house" might refer to the physical structure, while "home" often conveys emotional attachment and belonging.

• "Fast" and "quick" are commonly translated as "rápido" in Spanish, yet "fast" emphasizes speed, whereas "quick" implies rapidity or agility.

Crossref doi | Google Scholar | WorldCat | MENDELEY

2. Semantic Divergence: Synonyms may differ in meaning due to cultural, pragmatic, or lexical considerations. pragmatic associations for example

In lexical priming, Hoey (2005) points out: Just as a word or word sequence may be primed for semantic association, so it may be primed pragmatically as well. Pragmatic association occurs when a word or word sequence is associated with a set of features that all serve the same or similar pragmatic function (e.g. indicating vagueness, uncertainty). Parallel corpora reveal such nuances by comparing translations of synonyms in different contexts. For example, while "cheap" and "inexpensive" transmit comparable economic notions, their implications and usage varies, reflecting cultural preferences and socioeconomic settings. • "Brave" and "courageous" share the notion of bravery, but "brave" may connote fearlessness or boldness, while "courageous" emphasizes moral fortitude.

• "Childish" and "childlike" both relate to characteristics of a child, but "childish" often carries negative connotations of immaturity, while "childlike" suggests innocence or wonder.

3. Contextual Convergence: • In formal contexts like academic or legal writing, "begin" and "commence" can be used interchangeably to indicate the start of an event or process.

• "Pleasant" and "agreeable" may overlap in expressions of contentment or approval, indicating that both communicate a favorable attitude about an experience or contact.

• "Car" and "automobile" may converge in technical or formal contexts, where both terms are used to describe motor vehicles.

• "Laugh" and "chuckle" may converge in informal or casual contexts, where both denote forms of amusement or mirth, albeit with varying degrees of intensity.

CONCLUSION

In conclusion, the semantic analysis of synonyms across parallel corpora reveals complex linguistic nuances that have a considerable impact on natural language processing tasks. Researchers learn about how synonyms are employed across languages and circumstances by examining translation equivalence, semantic divergence, and contextual convergence. This understanding increases not only the accuracy of machine translation, but also sentiment analysis and information retrieval. By drawing on this information, NLP systems can better capture the nuances of human language, resulting in more effective communication and interaction between humans and machines. As parallel corpora continue to grow and adapt, more research into synonym semantics promises to increase

NLP skills and contribute to its larger applications in a variety of fields.

REFERENCES

1. Baroni, M., & Bernardini, S. (2004). BootCaT: Bootstrapping Corpora and Terms from the Web. In Proceedings of LREC.

2. Mikolov, T., et al. (2013). Distributed Representations of Words and Phrases and their Compositionality. In Proceedings of NIPS.

3. Resnik, P., & Diab, M. (2000). Selection and Information: A Class-Based Approach to Lexical Relationships.

4. In Proceedings of ACL. Goodman, N. (1952) 'On likeness of meaning', in Linsky, L. (ed.) Semantics and the philosophy of language: A collection of readings. Urbana, Illinois: University of Illinois Press, pp. 67-74.

5. Original work published (1949) Analysis 13, 1-7.

6. Gove, P. B. (ed.) (1984) Webster's New Dictionary of Synonyms. Springfield, MA: Merriam-Webster.

7. Gowers, E. (1986). The Complete Plain Words. 3 rd edn. London: H.M.S.O.