



THE SIGNIFICANCE OF SYNTACTIC ANALYSIS IN COMPUTATIONAL LINGUISTICS: ILLUSTRATED THROUGH RESEARCH ON WORD COMBINATIONS AND VALENCY

Nurmuxammedova Nurjaxon Istam kizi

Master's student at Samarkand State Institute of Foreign Languages

nurmuxammedovanurjaxon@gmail.com

Asadov Rustam Muminovich

Samarkand State Institute of Foreign Languages PhD, dotsent

rustamasadov1972@gmail.com

Annotation: This article explores the crucial role of syntactic analysis in computational linguistics, particularly focusing on how it aids in understanding word combinations and valency. By examining recent research and methodologies, we highlight the implications of syntactic structures for natural language processing (NLP) tasks, such as machine translation, sentiment analysis, and information retrieval. The findings underscore the necessity of robust syntactic frameworks to enhance the performance of NLP systems.

Keywords: Syntactic analysis, computational linguistics, word combinations, valency, natural language processing, machine translation, sentiment analysis.

Syntactic analysis is a foundational aspect of computational linguistics that deals with the structure of sentences and the relationships between words. It plays a pivotal role in understanding how words combine to form meaningful phrases and sentences. This significance is particularly evident in the study of word combinations and valency - two concepts that illuminate how words interact within a syntactic framework.

Syntactic analysis, also known as parsing, is the process of analyzing a string of symbols in terms of its grammatical structure. In the context of language processing, this involves examining sentences to determine their syntactic structure and ensuring that each word follows the linguistic rules.

In computer science, syntactic analysis is critical for converting high-level programming languages into machine code.

Word combinations refer to the ways in which words can be grouped together to convey specific meanings, while valency pertains to the number of arguments a verb can take. Together, these concepts form the basis for understanding sentence structure and meaning, which are essential for various applications in natural language processing (NLP). This article aims to elucidate the importance of syntactic analysis in computational linguistics by examining its implications for word combinations and valency [1].

Computational linguistics is an interdisciplinary field concerned with the computational modelling of natural language, as well as the study of appropriate computational approaches to linguistic questions. In general, computational linguistics draws upon linguistics, computer

science, artificial intelligence, mathematics, logic, philosophy, cognitive science, cognitive psychology, psycholinguistics, anthropology and neuroscience, among others. Computational linguistics is closely related to mathematical linguistics.

In computational linguistics, syntactic analysis uses algorithms to transform text into structured formats that computers can process. Parsing algorithms are essential for tasks such as code compilation in computer science, as well as natural language processing tasks like sentiment analysis. The context-free grammar is often utilized because it is powerful enough to describe most of the syntax used in programming languages. For example, most modern programming languages like Python, Java, and C++ use context-free grammars for their parser implementations. These grammars are defined using a set of recursive rules or productions that describe which strings of symbols comprise syntactically correct strings in the language.

Lexical valency is defined as the aptness of a word to appear in various combinations. The range of the lexical valency of words is linguistically restricted by the inner structure of the English word-stock. This can be easily observed in the selection of synonyms found in different word-groups. Though the verbs lift and raise are usually treated as synonyms, it is only the latter that is collocated with the noun question. The verb take may be synonymically interpreted as grasp, seize, catch, lay hold of, etc., but it is only take that is found in collocation with the nouns examination, measures, pre-cautions, etc.; only catch in catch smb. napping and grasp in grasp the truth [2. p. 64-65]. Lexical collocability is understood as the ability of a lexeme to collocate with a class of lexemes (allolexes) on the basis of the common classeme, which is a case of semantic accord. The class of lexemes forms the domain of the use of the lexeme delimited by the explanation of its sememe. This explanation represents the semantic norm of the lexeme [3. p. 176]. The terms collocability and valency are considered to be distinct in the following manner: the former is defined as a particular actualization of a lexical unit's ability to collocate with other lexical units, leading to a conceptual combination; the latter is opposed as potential ability of lexical units to collocate with other units of the same category [4. p. 75]. Valency is also understood as the ability of lexical units to interact with other units at close and distant relations, or in micro and macro contexts. Valency is related to such notions as 'position' and 'function'. According to S.D. Katznelson, in such languages as English valency and collocability are purely distinct [5. p. 20]

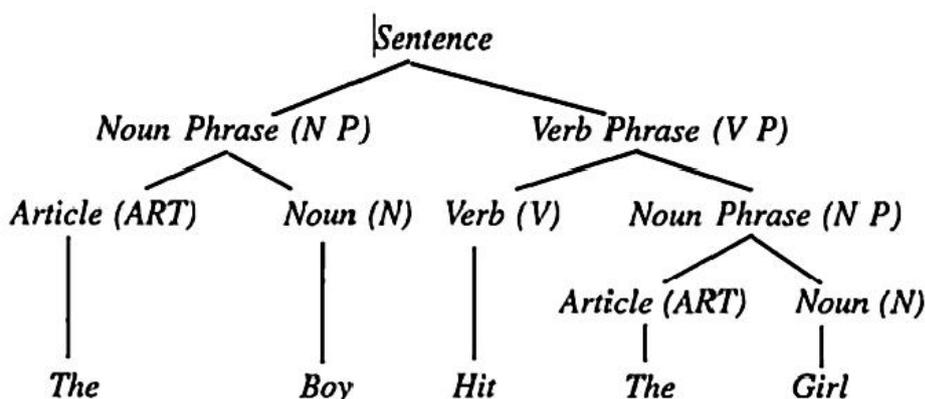


Fig. 1. Structural representation of a sentence /
Рис. 1. Структурная репрезентация предложения

Valency refers to the number and type of arguments a verb requires (e.g., subject, direct object, indirect object). For example, the verb “give” is typically ternary valency—it needs three arguments: a giver (subject), a thing given (direct object), and a recipient (indirect object).

Importance in NLP

Correctly identifying valency frames is crucial for parsing sentence meaning and for tasks like semantic role labeling. For instance, the verb “run” can have different valency frames:

- Intransitive: "She runs." (one argument)
- Transitive: "She runs a business." (two arguments)

Syntactic analysis enables algorithms to detect these frames, guiding accurate interpretation and generation of sentences.

Syntactic analysis involves parsing sentences to reveal their grammatical structure. It allows computational models to understand not only the individual words but also their relationships and functions within a sentence. This understanding is crucial for tasks such as:

1. Machine Translation. Accurate translation requires an understanding of the grammatical structures of both source and target languages. Syntactic analysis helps identify equivalent structures and ensures that meaning is preserved.
2. Sentiment Analysis. The sentiment expressed in a sentence often hinges on syntactic structures. For instance, negation can drastically alter sentiment; thus, syntactic analysis aids in correctly interpreting sentiment-bearing phrases.
3. Information Retrieval. Effective search engines utilize syntactic analysis to improve query understanding, ensuring that results are relevant to the user's intent.

Word combinations encompass collocations, idiomatic expressions, and other fixed or semi-fixed phrases that do not always follow standard grammatical rules. Understanding these combinations requires sophisticated syntactic models that can capture their unique properties.

Recent research has shown that statistical models, such as n-grams and neural networks, can effectively identify common word combinations. However, these models often lack the ability to account for nuanced syntactic relationships. Incorporating syntactic analysis enables better recognition of these combinations, leading to improved performance in NLP applications.

Example: Collocation Identification

Consider the phrase "strong coffee." A syntactic analysis reveals that "strong" modifies "coffee," indicating a specific type of coffee rather than any coffee. By recognizing such combinations, NLP systems can enhance their understanding of user queries or improve text generation capabilities.

Valency refers to the number of arguments a verb can take and is essential for constructing grammatically correct sentences. Understanding valency helps determine how verbs interact with nouns, pronouns, and other elements within a sentence.

Research has shown that different languages exhibit varying valency patterns, which must be accounted for in multilingual NLP applications. By employing syntactic analysis to model these patterns, researchers can develop more accurate language models that respect the grammatical rules of each language.

Example: Verb Argument Structure

In English, the verb "give" typically requires three arguments: a subject (the giver), an object (the gift), and an indirect object (the recipient). A syntactic analysis of the sentence "She gave

him a book" reveals this structure clearly. Failing to recognize such valency can lead to incorrect interpretations in automated systems.

Despite its importance, syntactic analysis faces several challenges. Ambiguity in language, variations in syntax across different languages, and the complexity of natural language structures pose significant hurdles. Future research should focus on developing more sophisticated models that integrate syntactic analysis with semantic understanding to address these challenges effectively.

Advancements in deep learning and neural network architectures offer promising avenues for improving syntactic parsing and understanding. By leveraging large datasets and contextual embeddings, researchers can create models that better capture the intricacies of syntax and semantics.

Conclusion

To sum up, syntactic analysis is a cornerstone of computational linguistics that significantly enhances our understanding of word combinations and valency. As demonstrated through various applications in natural language processing, robust syntactic frameworks are essential for improving machine translation, sentiment analysis, and information retrieval systems.

The ongoing research in this field highlights the need for integrating syntactic analysis with advanced computational techniques to tackle existing challenges. By doing so, we can enhance the performance of NLP systems, ultimately leading to more accurate and contextually aware applications. The future of computational linguistics lies in our ability to bridge the gap between syntax and semantics, paving the way for more intelligent and responsive language technologies.

References:

1. Kобрina, N.A. (2007). Is there any regular correspondence between the lexical meaning of a verb and its grammatical paradigm? *Issues of Cognitive Linguistics*, 4, 40-43. (In Russ.)
2. Ginzburg, R.S., Khidekel, S.S., Knyazeva, G.Y., & Sankin, A.A. (1979). *A course in modern English lexicology*. Moscow: Higher School. (In Russ.)
3. Filipec, J. (1994). *Lexicology and lexicography: Development and state of the research*. In: *Prague school of structural and functional linguistics*, Ph.A. Luelsdorff (Ed). Amsterdam: John Benjamins B.V. pp. 163-185.
4. Yudina, N.V. (2006). On some new phenomena in lexeme collocating in the modern Russian language (based on attributive-substantive complexes). *Izvestia: Herzen University Journal of Humanities and Science*, 16, 75-84. (In Russ.)
5. Katznelson, S.D. (1987). To the notion of types of valency. *Voprosy Jazykoznanija (Topics in the study of language)*, 3, 20-32.
6. Kalinina, O.N. (2008). Unity of the notions valency, function and position. *Izvestia: Herzen University Journal of Humanities and Science*, 69, 111-114.
7. Асадов Р.М. Синтаксическая валентность на примере синтаксического анализа трехвалентных элементов в позиции неядерного оппозитивного предикцирующего компонента (пар2) *Вестник Челябинского государственного...*, 2016. С. 25-35
8. Nurmuxammedova, N. (2024). The Importance of Syntactic Analysis in Computer Linguistics (in Research Examples of Word Combinations and Valency). *Journal of Language Pedagogy and Innovative Applied Linguistics*, 2(5), 74-77. <https://doi.org/10.1997/c8eghc44>
9. Karimova M., Asadov R., Aminova N., *Proficient English users vocabulary Dictionary for advanced and proficiency English learners 2020*. <http://www.morebooks.shop/>
10. Rustam Asadov, Shohista Mardiyeva, *Syntactic Valency of Predicative Components Conference Proceedings: Fostering Your Research.... 2024* C. 159-162.