PROBLEMS AND SOLUTIONS OF SCIENTIFIC AND INNOVATIVE RESEARCH
Volume 02, Issue 02, 2025

STRUCTURING OF MOTION AND STATE-EXPRESSING VERBS IN UZBEK LANGUAGE SYNSETS

Najmiddinov Muhammadjon
Doctor of Philosophy in Philology, Associate Professor, Kokand University.

Abstract: This thesis explores the structuring of motion and state-expressing verbs in Uzbek language synsets, emphasizing the importance of semantic accuracy and computational methodologies in linguistic research. It discusses the principles of verb synset formation, including synonym selection, contextual verification, and differentiation of polysemous verbs. The role of lexical databases such as WordNet and NLP-based techniques in automating synset development is examined. The findings highlight the need for a balance between computational efficiency and linguistic expertise to ensure precise semantic grouping, ultimately enhancing natural language processing applications and linguistic research.

Keywords: Uzbek language, verb synsets, lexical databases, computational linguistics, NLP, synonym selection, semantic analysis, WordNet, linguistic research

The structuring of motion and state-expressing verbs in Uzbek language synsets is an essential linguistic process that contributes to the development of lexical databases, such as WordNet, and enhances natural language processing (NLP) applications. A synset is a collection of words that share a common meaning, allowing for better organization of vocabulary and facilitating computational language understanding. In the case of verbs, synsets are formed by grouping verbs that express similar actions or states, ensuring that they maintain semantic coherence. The process of building verb synsets requires a detailed analysis of verb meanings, synonym selection, contextual verification, and differentiation of polysemous verbs. A verb's meaning must first be precisely defined using linguistic resources, including explanatory dictionaries and language corpora. This step ensures that each verb's core semantic components, or semes, are correctly identified. Identifying the primary meaning of a verb is especially crucial when dealing with polysemous verbs, as they may belong to multiple synsets depending on their usage in different contexts. For instance, a verb with both literal and figurative meanings might need separate synsets to preserve semantic clarity. Once a verb's meaning is established, potential synonyms must be identified. Synonym dictionaries, thesauri, and NLP-based semantic models can assist in this process by

suggesting verbs that share similar meanings. However, finding words with identical meanings is rare, making it necessary to evaluate the degree of synonymy. The contextual applicability of each synonym is assessed through linguistic corpora or explanatory dictionaries, ensuring that they can replace the base verb without altering the meaning significantly. The formation of a synset involves grouping the selected synonyms into a cohesive unit, ensuring that all included verbs share a fundamental meaning. For example, the verb "bormoq" (to go) can be associated with "qadam tashlamoq" (to step), "harakat qilmoq" (to move), "yo‘nalmoq" (to head), and "ergashmoq" (to follow). These verbs all imply motion but differ in nuances. If a verb exhibits multiple meanings, it may belong to multiple synsets, requiring careful differentiation of its semantic interpretations. For example, "o‘rganmoq" (to learn) might appear in one synset related to education and another related to adaptation or habituation. Contextual verification is essential to ensure that the verbs within a synset are genuinely interchangeable in sentences. Some verbs may appear synonymous in one context but diverge in another. For instance, while "bormoq" and "yo‘nalmoq" both indicate movement, "yo‘nalmoq" implies a more purposeful directionality, which may not always align with "bormoq." Distinguishing between subtle variations in meaning prevents semantic inaccuracies.

A further level of refinement involves distinguishing verb meanings based on semantic shades or nuances. If verbs within a synset exhibit slight variations in meaning, they may be separated into subgroups to enhance accuracy. The syntactic behavior of verbs also plays a role in synset formation, as verbs with different valency patterns may not always function as true synonyms. The final step in synset construction is verification and correction. Each synset undergoes a review to ensure that all included verbs align with the intended meaning. Any verbs that do not perfectly match the semantic scope are either removed or reassigned to more suitable synsets. This process guarantees that synsets maintain semantic consistency and practical applicability. Reliable linguistic resources are essential for constructing accurate synsets. WordNet, for instance, serves as a foundational linguistic database that organizes words into semantic groups. It provides structured information on verbs, including their synonym relationships and hierarchical associations. Thesauri and synonym dictionaries also serve as valuable tools for finding semantically related words. Additionally, corpus analysis allows for real-world verification of verb meanings, helping to refine synset structures. The principles underlying synset development emphasize semantic proximity, contextual interchangeability, and

meaning differentiation. Verbs within a synset should exhibit a high degree of semantic similarity, ideally allowing for mutual substitution without altering the sentence's fundamental meaning.

However, complete interchangeability is rare, and thus, contextual factors must always be considered. The stylistic and connotative properties of verbs must also be taken into account. Even if two verbs share a core meaning, differences in emotional tone or stylistic register may affect their suitability within a synset. For example, a neutral verb like "gapirmoq" (to speak) may not fully align with an informal variant like "g‘o‘ldiramoq" (to mumble). Considering such distinctions ensures that synsets remain precise and functional. The development of verb synsets also involves hierarchical relationships such as hypernymy and hyponymy. Hypernyms represent broader categories, while hyponyms denote specific instances. For example, "harakat qilmoq" (to move) is a hypernym encompassing more specific verbs like "yugurmoq" (to run) and "sakramoq" (to jump). Recognizing these relationships allows for a more structured organization of verb synsets. Another relevant concept is troponymy, which describes how one verb specifies a manner of performing another verb’s action. For example, "shivirlamoq" (to whisper) is a troponym of "gapirmoq" (to speak) because it specifies a particular way of speaking. Differentiating troponyms from true synonyms ensures that synsets remain semantically accurate. Computational methods, including NLP techniques and machine learning, have been increasingly utilized in synset development. Automatic methods involve analyzing large text corpora to detect word co-occurrence patterns and semantic similarities. Distributional semantics models, such as word embeddings, help identify related verbs by examining their contextual usage. Machine learning approaches, including deep learning models like Word2Vec and BERT, can further refine synset construction by recognizing complex semantic relationships. However, despite advancements in automation, expert linguistic validation remains necessary to ensure accuracy. The process of constructing verb synsets must balance linguistic intuition with computational efficiency. While automated methods can process vast amounts of data, human expertise is crucial for resolving ambiguities and refining semantic groupings. The iterative nature of synset development involves continuous refinement based on linguistic insights and evolving language usage. Properly constructed verb synsets enhance various linguistic applications, including machine translation, information retrieval, and language learning. By providing structured representations of verb meanings, synsets contribute to more accurate text analysis and semantic processing. They also

support lexical databases and thesauri, facilitating word sense disambiguation and improving computational models of language. In conclusion, the development of verb synsets in the Uzbek language is a meticulous process that involves defining meanings, selecting synonyms, verifying contextual applicability, and refining semantic structures. The principles of semantic proximity, contextual consistency, and hierarchical relationships guide the organization of verb synsets. Computational techniques, while valuable, must be complemented by expert linguistic analysis to ensure accuracy. Ultimately, well-structured verb synsets contribute to linguistic research and NLP advancements, enhancing language understanding and computational processing.

References:

1. Fellbaum, C. (1998). WordNet: An electronic lexical database. MIT Press.
2. Miller, G. A. (1995). WordNet: A lexical database for English. Communications of the ACM, 38(11), 39-41.
3. Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation, and application of a wide-coverage multilingual semantic network. Artificial Intelligence, 193, 217-250.
4. Pustejovsky, J. (1995). The generative lexicon. MIT Press.