

INTEGRATING CORPUS LINGUISTICS INTO TEACHING ENGLISH POLYSEMY: AN EXPERIMENTAL STUDY WITH TEFL LEARNERS

Barnoyeva Nilufar Yoqubovna

Associate professor, Department of Foreign Languages

Tashkent University of Information Technologies

named after Muhammad al- Khwarizmi

barnoyevanilufar1984@gmail.com

Tel: 977763058

Yangiyeva Mohinur Asror qizi

“Tashkent Institute of Irrigation and

Agricultural Mechanization Engineers”

National Research University

Abstract: Polysemy, the phenomenon whereby a single lexical item possesses multiple related meanings, represents one of the most challenging aspects of vocabulary acquisition for learners of Teaching English as a Foreign Language (TEFL). Traditional vocabulary instruction often presents words in isolation and focuses primarily on their most common meanings, limiting learners' ability to comprehend and use polysemous words accurately in different contexts. Corpus Linguistics offers an innovative solution by providing authentic language data that enables learners to observe how words function across diverse communicative situations. This study investigates the effectiveness of integrating Corpus Linguistics into teaching English polysemy. An experimental design involving university TEFL students was employed to compare corpus-based instruction with traditional vocabulary teaching methods. The findings indicate that learners exposed to corpus-based activities demonstrated significantly higher levels of polysemy awareness, contextual interpretation, and productive vocabulary use. The study highlights the pedagogical value of Corpus Linguistics in facilitating deeper lexical understanding and recommends its integration into EFL vocabulary instruction.

Keywords: Corpus Linguistics, Polysemy, TEFL, Data-Driven Learning, Vocabulary Acquisition, Corpus-Based Teaching.

Annotatsiya: Polisemiya, ya'ni bitta leksik birlikning o'zaro bog'liq bo'lgan bir nechta ma'nolarga ega bo'lish hodisasi, ingliz tilini chet tili sifatida o'qitish (TEFL) yo'nalishidagi talabalar uchun lug'at boyligini o'zlashtirishning eng murakkab jihatlaridan biri hisoblanadi. An'anaviy lug'at o'qitish usullari ko'pincha so'zlarni alohida holda taqdim etib, asosan ularning eng keng tarqalgan ma'nolariga e'tibor qaratadi. Bu esa o'quvchilarning polisemantik so'zlarni turli kontekstlarda to'g'ri tushunish va qo'llash imkoniyatlarini cheklaydi. Korpus lingvistikasi o'quvchilarga so'zlarning turli kommunikativ vaziyatlarda qanday qo'llanishini kuzatish imkonini beruvchi autentik til ma'lumotlarini taqdim etish orqali innovatsion yechim taklif etadi.

Mazkur tadqiqot korpus lingvistikasini ingliz tilidagi polisemiyaning o'qitish jarayoniga integratsiya qilish samaradorligini o'rganadi. Korpusga asoslangan ta'limni an'anaviy lug'at o'qitish usullari bilan taqqoslash maqsadida universitetning TEFL yo'nalishidagi talabalari ishtirokida eksperimental tadqiqot o'tkazildi. Natijalar korpusga asoslangan faoliyatlarda qatnashgan talabalar polisemiya bo'yicha xabardorlik, kontekstual talqin qilish hamda so'z boyligidan produktiv foydalanish ko'rsatkichlari bo'yicha sezilarli darajada yuqori natijalarga erishganligini ko'rsatdi. Tadqiqot korpus lingvistikasining leksik birliklarni chuqurroq anglashni rivojlantirishdagi pedagogik ahamiyatini yoritadi hamda uni ingliz tilini chet tili sifatida o'qitish jarayonidagi lug'at ta'limiga integratsiya qilishni tavsiya etadi.

Kalit so'zlar: Til korpusi, polisemiya, ingliz tilini chet tili sifatida o'qitish, ma'lumotlarga asoslangan ta'lim (Data-Driven Learning), lug'at boyligini o'zlashtirish, korpusga asoslangan o'qitish.

Аннотация: Полисемия, то есть явление, при котором одна лексическая единица обладает несколькими взаимосвязанными значениями, представляет собой один из наиболее сложных аспектов усвоения словарного запаса для студентов направления преподавания английского языка как иностранного (TEFL). Традиционные методы обучения лексике часто рассматривают слова изолированно и сосредотачиваются главным образом на их наиболее распространённых значениях, что ограничивает способность обучающихся правильно понимать и использовать многозначные слова в различных контекстах. Корпусная лингвистика предлагает инновационное решение, предоставляя аутентичные языковые данные, которые позволяют обучающимся наблюдать за тем, как слова функционируют в различных коммуникативных ситуациях. В данном исследовании рассматривается эффективность интеграции корпусной лингвистики в процесс обучения английской полисемии. Для сравнения корпусно-ориентированного обучения с традиционными методами преподавания лексики был использован экспериментальный дизайн с участием студентов университета, обучающихся по направлению TEFL. Полученные результаты показали, что студенты, участвовавшие в корпусно-ориентированных учебных заданиях, продемонстрировали значительно более высокий уровень осведомлённости о полисемии, навыков контекстуальной интерпретации и продуктивного использования лексики. Исследование подчёркивает педагогическую ценность корпусной лингвистики в формировании более глубокого понимания лексических единиц и рекомендует её интеграцию в обучение лексике английского языка как иностранного.

Ключевые слова: корпусная лингвистика, полисемия, TEFL, обучение на основе данных (Data-Driven Learning), усвоение словарного запаса, корпусно-ориентированное обучение.

Vocabulary knowledge constitutes a fundamental component of communicative competence in second language learning. Among various lexical phenomena, polysemy presents particular difficulties for TEFL learners because a single word may express several related meanings depending on context. For example, the word head may refer to a part of the body, a leader of an organization, the top of an object, or a source of water. Such semantic variation often causes misunderstanding and inappropriate language use among learners.

Traditional teaching approaches typically introduce vocabulary through dictionary definitions, translation, or memorization of isolated meanings. While these methods may facilitate initial vocabulary recognition, they often fail to develop learners' ability to interpret and employ polysemous words in authentic contexts. Recent developments in Corpus Linguistics have created opportunities for more effective vocabulary instruction. Through access to large collections of authentic texts, learners can observe multiple meanings of a word, identify patterns of use, and discover semantic relationships independently. This study explores the integration of Corpus Linguistics into teaching English polysemy and evaluates its impact on EFL learners' vocabulary development.

Polysemy refers to the coexistence of several related meanings within a single lexical item. Unlike homonymy, where meanings are unrelated, polysemous meanings share semantic connections. Scholars have emphasized that mastering polysemy is essential for achieving advanced language proficiency because polysemous words occur frequently in authentic communication.

Research indicates that TEFL learners often rely on the primary meaning of a word and experience difficulties recognizing less frequent meanings. Consequently, learners may misunderstand texts or produce inaccurate utterances when encountering unfamiliar semantic extensions.

Corpus Linguistics involves the analysis of naturally occurring language through computerized corpora. Since the emergence of corpus-based pedagogy, researchers have highlighted its potential to improve vocabulary learning by exposing students to authentic language patterns.

Data-Driven Learning (DDL), introduced by Johns (1991), encourages learners to act as language researchers by examining concordance lines and discovering linguistic regularities independently. Corpus-based approaches have been successfully applied to teaching grammar, collocations, phraseology, and vocabulary. However, comparatively few studies have focused specifically on polysemy instruction.

Corpus tools enable learners to observe multiple uses of a word across different contexts. Through concordance analysis, students can identify recurring semantic patterns and distinguish between literal and metaphorical meanings. Such exposure promotes contextualized vocabulary learning and strengthens semantic awareness.

The Corpus of Contemporary American English (COCA) provides a particularly valuable resource for teaching polysemy because it contains authentic examples from spoken language, fiction, magazines, newspapers, blogs, web pages, and academic texts. Through concordance lines, learners can explore how a single lexical item acquires different meanings in different communicative contexts. For example, the word *head* exhibits several related meanings. In one context, it refers to a body part (She nodded her head and smiled). In another, it denotes a leader (He became the head of the department). The same word may also indicate the top position of an object (The head of the table was reserved for guests) or the source of a river (The river's head is located in the mountains). By examining multiple concordance lines, learners can recognize the semantic connections among these meanings.

Similarly, the verb *run* demonstrates extensive polysemy. It may describe physical movement (The children ran across the playground), operation of a machine or software (The software runs smoothly on most computers), management of an organization (She runs a successful online business), or continuation over time (The meeting ran for three hours). Corpus evidence enables learners to understand how contextual cues determine the intended meaning.

Another example is the word *issue*, which may refer to a problem or topic (Climate change is an important issue worldwide), a published edition (The latest issue of the journal was published in May), or the act of distribution (The bank will issue new credit cards next month).

For students of Information and Communication Technologies (ICT), discipline-specific polysemous vocabulary is particularly important. The word *network*, for instance, may refer to a computer system (The company upgraded its wireless network), a group of people or organizations (She built a professional network during her studies), or a broadcasting organization (The television network announced a new program). Exposure to authentic corpus examples helps learners understand how professional terminology extends beyond specialized contexts.

Researchers have reported that corpus-based activities enhance learners' inferencing skills, critical thinking abilities, and lexical retention. By examining authentic examples and identifying semantic patterns, learners develop a more nuanced understanding of word meanings than through traditional vocabulary instruction alone. Consequently, corpus-based approaches provide an effective pedagogical framework for teaching polysemy in EFL contexts.

The intervention lasted eight weeks. Students in the experimental group participated in corpus-based activities using the Corpus of Contemporary American English (COCA), while students in the control group received traditional vocabulary instruction.

Experimental Group. Students were introduced to corpus-search techniques and

analyzed concordance lines of highly frequent polysemous words, including head, run, issue, point, case, and network. The instructional sequence consisted of the following stages:

Observation of Concordance Lines

Students examined authentic examples extracted from COCA and identified contextual clues that determined word meaning.

Meaning Classification

Learners grouped concordance lines according to different meanings. For example, occurrences of run were categorized as physical movement, operation of software, business management, or duration.

Semantic Mapping

Students created visual semantic networks showing relationships among the various meanings of polysemous words.

Contextual Interpretation Tasks

Learners inferred meanings from unfamiliar corpus examples and justified their interpretations based on contextual evidence.

Productive Activities

Students wrote sentences and short paragraphs using different meanings of target words. For example, they were asked to use the word network in both technological and social contexts.

A typical classroom activity involved investigating the word point. Students analyzed approximately twenty concordance lines from COCA and classified them into categories such as argument or idea (That's an interesting point), location (the southernmost point of the island), purpose (What is the point of this exercise?), and score in a game (The team scored three points). Following the classification task, students discussed semantic relationships among the meanings and created semantic maps.

Control Group. Students learned the same target vocabulary through conventional methods, including dictionary definitions, translation exercises, memorization activities, and vocabulary drills. Instruction focused primarily on providing explanations and practice exercises rather than independent discovery of meaning patterns.

The results support previous research emphasizing the effectiveness of Corpus Linguistics in vocabulary instruction. Exposure to authentic corpus data enabled learners to observe how meanings change according to context and communicative purpose. Rather than memorizing isolated definitions, students developed deeper lexical knowledge through discovery-based learning.

The findings also confirm the pedagogical benefits of Data-Driven Learning. By analyzing concordance lines independently, learners became active participants in the

learning process. This autonomy contributed to greater engagement and improved retention of polysemous meanings.

Furthermore, corpus-based instruction promoted critical language awareness by encouraging students to examine patterns of use and semantic extensions. Such skills are particularly valuable for advanced TEFL learners who must interpret complex academic and professional texts.

This study investigated the integration of Corpus Linguistics into teaching English polysemy and demonstrated its effectiveness in enhancing learners' vocabulary development. Corpus-based instruction improved students' polysemy awareness, contextual interpretation, and productive vocabulary use more effectively than traditional teaching methods.

The findings suggest that Corpus Linguistics should be incorporated into TEFL vocabulary instruction, particularly when teaching polysemous words. Future research may explore the long-term effects of corpus-based learning and its application across different proficiency levels and teaching materials.

References:

1. Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. *English Language Research Journal*, 4, 1–16.
2. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
3. McCarthy, M., & O'Dell, F. (2017). *English Vocabulary in Use: Advanced*. Cambridge University Press.
4. Nation, I. S. P. (2013). *Learning Vocabulary in Another Language*. Cambridge University Press.
5. Sinclair, J. (2004). *Trust the Text: Language, Corpus and Discourse*. Routledge.
6. Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell Publishing.