

**DIFFERENTIATING BETWEEN FREE AND SET EXPECTATIONS:
CHALLENGES AND DISTINCTIVE FEATURES**

Author: Ubaydullayeva Madinabonu Lutfullo qizi
Scientific advisor: Mukhammadiyev Aziz Shavkatovich
Scientific advisor: Bafoeva Nilufar Salimovna

***Abstract:** linguists and programmers often encounter the distinction between free and set expressions, which can be a source of confusion and complexity in language analysis and natural language processing tasks. This article aims to provide a comprehensive overview of the problems associated with differentiating these two types of expressions and their distinctive features. We explore the challenges in identifying and categorizing free and set expressions, discuss the implications for language understanding and generation, and highlight the importance of addressing these issues in various applications, such as machine translation, text summarization, and information extraction. By understanding the nuances between free and set expressions, researchers and practitioners can develop more robust and accurate language processing systems.*

***Аннотация:** лингвисты и программисты часто сталкиваются с различием между свободными и устойчивыми выражениями, которое может стать источником путаницы и сложности при анализе языка и задачах обработки естественного языка. Эта статья призвана предоставить всесторонний обзор проблем, связанных с различием этих двух типов выражений, и их отличительных особенностей. Мы исследуем проблемы идентификации и категоризации свободных и устойчивых выражений, обсуждаем последствия для понимания и генерации языка, а также подчеркиваем важность решения этих проблем в различных приложениях, таких как машинный перевод, автоматическое реферирование и извлечение информации. Понимая нюансы между свободными и*

www.pedagoglar.org

устойчивыми выражениями, исследователи и практики могут разрабатывать более надежные и точные системы обработки языка.

***Annotatsiya:** tilshunoslar va dasturchilar ko'pincha erkin va o'rnatilgan iboralar o'rtasidagi farqni ko'radilar, bu til tahlili va tabiiy til protsessori vazifalarida chalkashlik va murakkablikka sabab bo'lishi mumkin. Ushbu maqola ushbu ikki turdagi iboralarni farqlash bilan bog'liq muammolar va ularning farqli xususiyatlarini atroflicha ko'rib chiqishni maqsad qiladi. Biz erkin va o'rnatilgan iboralarni aniqlash va turkumlash bilan bog'liq muammolarni o'rganamiz, til tushunish va generatsiyasiga bo'lgan oqibatlarini muhokama qilamiz va mashinaviy tarjima, matn umumlashtirish va axborot chiqarish kabi turli xil qo'llanmalarda ushbu muammolarni hal qilish muhimligini ta'kidlaymiz. Erkin va o'rnatilgan iboralar o'rtasidagi nozikliklarga tushunish orqali, tadqiqotchilar va amaliyotchilar yanada barqaror va aniq til jarayonlash tizimlarini yarata olishadi.*

***Key words:** free expressions, set expressions, language analysis, natural language processing (NLP), challenges in identification and categorization, implications for language understanding and generation, applications (machine translation, text summarization, information extraction), distinctive features, robust and accurate language processing systems*

***Ключевые слова:** свободные выражения, устойчивые выражения, анализ языка, обработка естественного языка (NLP), проблемы идентификации и категоризации, последствия для понимания и генерации языка, приложения (машинный перевод, автоматическое реферирование, извлечение информации), отличительные особенности, надежные и точные системы обработки языка.*

***Kalit so'zlar:** erkin iboralar, o'rnatilgan iboralar, til tahlili, tabiiy til protsessori (NLP), identifikatsiya va turkumlash muammolari, til tushunish va*
www.pedagoglar.org

generatsiyasi uchun oqibatlar, qo'llanmalar (mashinaviy tarjima, matn umumlashtirish, axborot chiqarish), farqli xususiyatlar, barqaror va aniq til jarayonlash tizimlar.

Introduction

In the realm of language analysis and natural language processing (NLP), the distinction between free and set expressions is a crucial yet often overlooked aspect.

Free expressions are those that can be constructed freely, with a high degree of flexibility in their composition and meaning. In contrast, set expressions are relatively fixed and idiomatic, with a specific meaning that cannot be easily derived from the individual components.

Differentiating between these two types of expressions presents several challenges, as the boundaries can be blurred, and certain expressions may exhibit characteristics of both.

This article delves into the problems associated with this differentiation and explores the distinctive features that characterize free and set expressions.

Challenges in Differentiating Free and Set Expressions

One of the primary challenges in differentiating free and set expressions lies in their linguistic and structural properties. Free expressions are composed of words that can be freely combined and rearranged according to the rules of grammar, while set expressions are fixed, idiomatic phrases that cannot be easily modified without altering their meaning [6].

However, the boundaries between these two categories are often blurred, as some expressions exhibit characteristics of both free and set forms [13]. This overlap can make it difficult to consistently classify certain linguistic constructions.

Another challenge arises from the fact that the distinction between free and set expressions is not always clear-cut, as some expressions can exhibit a degree of flexibility or variability [5]. For example, the idiomatic expression "let the cat out of the bag" can be modified to "let the cat out of the sack" without significantly altering its meaning. This flexibility can make it difficult to determine whether an

expression should be classified as free or set.

Furthermore, the recognition and classification of free and set expressions can be influenced by factors such as language proficiency, cultural context, and individual experiences [2].

Learners of a language, for instance, may have different perceptions of what constitutes a set expression compared to native speakers, leading to potential misunderstandings or classification errors.

Overall, the challenges in differentiating free and set expressions highlight the need for a more nuanced understanding of language and the development of more sophisticated techniques for identifying and categorizing these linguistic constructions, particularly in the context of natural language processing and other language-related applications.

Implications and Applications

The distinction between free and set expressions has significant implications for language understanding and generation in various applications. In machine translation, for example, the accurate identification and translation of set expressions is crucial, as literal translation often fails to convey the intended meaning[8].

Failure to recognize set expressions can lead to awkward or nonsensical translations, underscoring the importance of addressing this challenge in machine translation systems.

Similarly, in text summarization, the ability to distinguish free and set expressions can aid in the identification of key concepts and the generation of more coherent and meaningful summaries. [7].

Set expressions often convey idiomatic meanings that are crucial for understanding the overall context and significance of a text, and their proper handling can enhance the quality of summarization.

In information extraction tasks, the ability to recognize and accurately categorize free and set expressions can improve the accuracy of entity and relation extraction, as well as the identification of key events and facts within a given text [4]. Misclassifying these linguistic constructions can lead to errors in information

retrieval and analysis, underscoring the importance of addressing this challenge.

Conclusion

The differentiation between free and set expressions is a complex and nuanced aspect of language analysis, with significant implications for various language processing tasks.

By understanding the challenges and distinctive features of these expressions, researchers and practitioners can develop more robust and accurate language processing systems, leading to improved performance in applications such as machine translation, text summarization, and information extraction.

BIBLIOGRAPHY

1. Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.]
2. Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Palgrave Macmillan
3. Chakraborty, T., & Bandyopadhyay, S. (2010). Identification of reduplication in morphologically rich Indian languages. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)* (pp. 199-207).
4. Finlayson, M. A., & Kulkarni, N. (2011). Detecting and classifying mentions of novel with a postprocessor ensemble. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 92-100).
5. Jackendoff, R. (1997). Twistin' the night away. *Language*, 73(3), 534-559.
6. Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language*, 70(3), 491-538.
7. Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford University Press.

8. Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. In Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2002) (pp. 1-15).

9. Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics*, 36(5), 549-569.]Rethinking idiomaticity: A usage-based approach. A&C Black.

10. Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics*, 36(5), 549-569.]

11. Tsvetkov, Y., & Wintner, S. (2014). Identification of multiword expressions by combining multiple linguistic information sources. *Computational Linguistics*, 40(2), 449-468.

12. Wulff, S. (2008). *Rethinking idiomaticity: A usage-based approach*. A&C Black.