DISSERTATION ABSTRACT

Name: Alona Kononenko-Szoszkiewicz

Title: "Badanie wzorców (mor)fonotaktycznych: studium porównawcze wybranych języków germańskich i słowiańskich"

Translation: "Exploring (mor)phonotactic patterns: A comparative study of selected Germanic and Slavic languages"

Abstract:

This PhD thesis aimed to investigate and analyze the phonotactic and morphonotactic patterns which occur in German, English, Polish, Ukrainian and Russian. In order to examine the consonant clusters within these languages, I adopted an interdisciplinary perspective by utilizing corpus data as the primary source to explore phonotactic preferences. The utilization of written corpus data allowed to investigate variety, language-specific constraints, and frequency of occurrence of consonant clusters in different word positions of the studied languages. This crucial information was then applied to draw comparative generalizations.

Furthermore, to fulfill the objectives of this thesis, a comprehensive bibliometric analysis was carried out on the Scopus database. This analysis provides a comprehensive overview of the current landscape of linguistic studies utilizing corpus linguistic methods to examine phonetic and phonological phenomena. In this context, I discuss the prominent trends and prevailing types of corpora used in corpus linguistic analysis (written vs. spoken), while also providing information on the dominant languages explored in such studies. The primary objective of the bibliometric analysis was to determine the position of phonetic and phonological research within the realm of corpus linguistics. The findings of the bibliometric analysis highlighted a substantial number of studies conducted in the past decade, particularly focusing on the study of sounds within the domains of natural language processing, language perception and acquisition, and machine learning.

Overall, my doctoral thesis is a compilation of four thematically related research articles which investigate and compare phonotactic and morphonotactic patterns of German, Russian and Ukrainian with a reference to the existing research within this domain, primarily on Polish and English (Zydorowicz et. al. 2016). The choice of the languages is motivated by the fact that they belong to the different language families and present significant distinctions with regard to phonological and morphological patterns. All four articles follow the same methodology for the exploration of phonotactic preferences of consonant clusters and apply the

Net Auditory Distance principle (Dziubalska-Kołaczyk, 2014). The analysis of consonant clusters, categorized as either phonotactic or morphonotactic, raises various questions, few of which were addressed in the research articles included in the doctoral thesis. Namely, what is the type and token frequency of specific consonant clusters present in the corpora of the selected languages, what is their preference status according to the NAD, and whether there is a relationship between cluster preference and its frequency.

Research Article 1 serves as an introductory and explanatory study that explores the emergence of morphonotactic consonant clusters from a diachronic standpoint. By analysing data from selected Slavic, Germanic, Baltic, and Romance languages and others, we aimed to identify major historical processes that contributed to the evolution of morphonotactic consonant clusters. The study concludes that these processes involve the following: vowel loss, Indo-European ablaut, affixation, compounding, metathesis, final and consonant epenthesis. Moreover, we concluded that the majority of morphonotactic clusters could be categorized into two main types: phonologically derived clusters resulting from vowel loss, as observed in Slavic languages, and morphologically derived clusters resulting from concatenation, as observed in Germanic languages.

The subsequent two articles, Research Article 2 and Research Article 3, provided a quantitative analysis of German consonant clusters utilizing data from the Austrian Media Corpus. These articles provided a quantitative characterization of German patterns of consonantal morphonotactics and phonotactics from various perspectives, including phonological, morphological, typological, and corpus linguistic. Specifically, word-initial and word-final consonant clusters are analysed within the framework of the Beats-and-Binding Model (Dziubalska-Kołaczyk 2002). German data was compared to the Russian data which allowed to arrive at first tentative generalisations regarding typological differences. One of the main findings of these studies is that the claim that phonotactic clusters are more preferred than morphonotactic clusters has been disproven. However, the analysis of the Ukrainian as presented in Research Article 4 confirmed a general presumption that morphonotactic clusters tend to be marked and therefore dispreferred. Furthermore, no significant correlation was observed between cluster preference and its frequency across all languages examined. However, based on the analysis of five languages, it can be inferred that the longer cluster is, the more likely it is to be morphonotactic.

To conclude, the findings presented in this doctoral thesis provide insights into the study of phonotactic and morphonotactic patterns by application of cross-linguistic methods of

research. By adopting a comparative approach across different languages, this research expands our understanding of the underlying principles and mechanisms that govern phonotactics and morphonotactics.