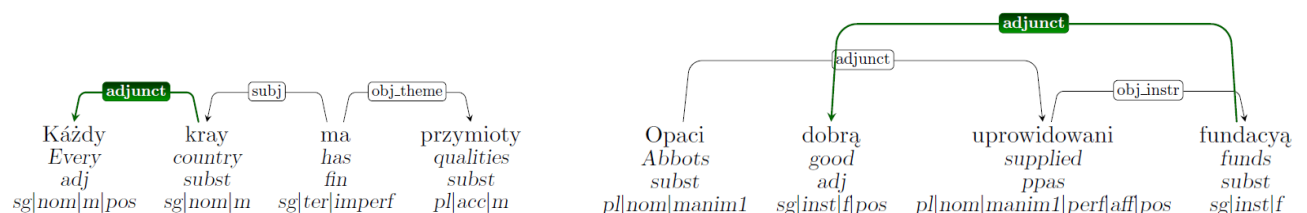


**Discontinuous noun phrases containing adjective or adjective-like modifiers in Middle Polish texts.
Preliminary research conducted on an experimental dependency treebank**

In my presentation, I will show examples of unusual – compared to the modern state – order of words in noun phrases containing an adjective or adjective-like modifier in the Polish language of the 17th and 18th centuries (e.g. *dziwna trafiła się awantura* ‘there was a **strange** riot’, lit. ‘**strange** there was a **riot**’; o *pożytkach z xiąg tłumaczenia **wynikających*** ‘about the **benefits** resulting from translation of books’, lit. ‘about the **benefits** from books’ translation **resulting**’; u *meگو widział mię **stryia*** ‘he saw me at **my** uncle’s’, lit. ‘at **my** he saw me **uncle’s**’). The uniqueness of the quoted and similar examples from Middle Polish texts lies in the fact that other sentence elements are relatively often placed between the components of nominal phrases, especially those elements that are not components of this nominal phrase. The elements of such discontinuous constructions can be so far apart as to make the sentence difficult to understand. Contrary to the current state of the language, such an order was very common in Middle Polish (cf. e.g. Ostaszewska 2002).

This phenomenon is known to historians of the Polish language, but the novelty of my research is that it is conducted on a dependency treebank (a syntactically annotated corpus) – the first such resource created for pre-modern Polish. Middle Polish Micro-Treebank contains 1,000 sentences selected from the Electronic Corpus of the 17th- and 18th-century Polish Texts (KorBa, www.korba.edu.pl). The treebank was created as an experimental resource; a further expansion is planned. One of the aims of the study presented here was to show what possibilities of studying historical syntax are offered by such a resource.

A dependency tree is a graph that unambiguously reflects the dependencies between the components of a sentence. It reflects e.g. the left or right position of the dependent relative to its governor (e.g. *dziwna awantura* vs. *awantura dziwna*) and the distance between them. It is worth mentioning that the graphical visualization of trees shows certain features of word order. In the case of a “typical”, linear word order, graph edges (visualized as arrows) do not intersect (as in the phrase *Każdy kraj ma przymioty* ‘Each country has qualities’). In the case of inversion and discontinuity of noun phrases with an adjective, we get trees with crossing edges (*Opaci dobrą uprowidowani fundacją* ‘Abbots supplied with good funds’, lit. ‘Abbots good supplied funds’).



Such a syntactically annotated corpus allows for various studies on the relative position of a noun and its modifier(s). In my speech, I will present some possible analyses and their results, carried out on a sample of Middle Polish texts.

References (selection)

Berdicevskis, A., and Eckhoff, H. (2020). *A Diachronic Treebank of Russian Spanning More Than a Thousand Years*. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pp. 5251–5256, Marseille, France. European Language Resources Association.

<https://aclanthology.org/2020.lrec-1.646/>

Gruszczyński, W., Adamiec, D., Bronikowska, R., Kieraś, W., Modrzejewski, E., Wieczorek, A., and Woliński, M. (2022). *The Electronic Corpus of 17th- and 18th-century Polish Texts*. *Language Resources and Evaluation* 56, 309–332.

Klemensiewicz, Z. (1968). *Zarys składni polskiej [An outline of the Polish syntax]*. Warszawa: PWN.

Mel'čuk, I. A. (1988). *Dependency syntax: theory and practice*. Albany: SUNY Press.

Ostaszewska, D. (Ed.) (2002). *Polszczyzna XVII wieku. Stan i przeobrażenia [Polish language of the 17th century. State and transformations]*, Katowice.

Wróblewska, A., and Rybak, P. (2019). Dependency parsing of Polish. *Poznań Studies in Contemporary Linguistics*, 55(2):305–337.