

Magdalena Derwojedowa, Witold Kieraś, Danuta Skowrońska

Instytut Języka Polskiego

Uniwersytet Warszawski

ul. Krakowskie Przedmieście 26/28

00-927 Warszawa

tel.: 22 552 03 97

e-mail: derwojed@uw.edu.pl

w.kieras@uw.edu.pl

danuta.skowronska@gmail.com

Poland

Robert Wołosz

Wydział Filologii Słowiańskiej

PTE BTK

Szláv Filológia Tanszék

Pécs, Ifjúság útja 6.

7624 Węgry

tel.: +36 72 503-600 w. 24138

e-mail: wolosz.robert@pte.hu

Hungary

PF 2014 (LXV): 251–256

Korpus polszczyzny xix wieku – od mikrokorpusu do korpusu średniej wielkości

A corpus of the 19th-century Polish – from a microcorpus to a medium-sized corpus

Słowa kluczowe: korpus diachroniczny; XIX wiek; analiza morfologiczna; ewolucja języka

Keywords: diachronic corpus; 19th century; morphological analysis; language evolution

Summary

The paper presents a general idea of a corpus of the 19th-century Polish texts. The project aims at bridging the gap between the corpora of earlier and further stages of language evolution, namely the corpus of Polish of the 17th-mid-18th century (under development) and the National Corpus of Polish. We believe that a small corpus of texts from the second half of the 19th century, that is currently being compiled with the aim to test a morphological analyzer, can be a germ of a larger corpus. In the paper we discuss some basic issues, such as principles of the texts selection and designing tools (or rather introducing modifications to the tools already developed).