



ANDRZEJ CHARCIAREK

 ORCID: <https://orcid.org/0000-0001-5651-5730>

Uniwersytet Śląski w Katowicach

## WSTĘP

Językoznawstwo jako dyscyplina nie zamknęło się w ujęciach strukturalistycznych czy formalistycznych i wciąż poszukuje nowych perspektyw badawczych. Początkowe skupienie uwagi na zagadnieniach systemowo-formalnych z czasem ustąpiło miejsca optyce poszerzonej o czynniki pozasystemowe. Stąd pojawienie się takich dziedzin w obrębie badań językoznawczych, jak socjolingwistyka, psycholingwistyka, pragmalingwistyka, etnolingwistyka czy mediolingwistyka. Wszystkie one wniosły i wciąż wnoszą wiele do opisu funkcjonowania języka w tekstach w określonym kontekście społeczno-kulturowym.

Nie inaczej jest z językoznawstwem korpusowym, które na dobre rozwinęło się w latach osiemdziesiątych ubiegłego wieku. Jego początki datują się od końca lat sześćdziesiątych i są związane z powstaniem pierwszego korpusu komputerowego zawierającego teksty w amerykańskiej odmianie angielszczyzny. Nie dziwi zatem fakt, że amerykańskie i angielskie środowiska naukowe miały największe zasługi dla rozwoju zarówno metodologii językoznawstwa korpusowego, jak i zasad tworzenia korpusów tekstów.

Początki narodowych korpusów języków słowiańskich to już lata dziewięćdziesiąte XX i pierwsza dekada XXI wieku. Wtedy to powstają narodowe korpusy języka czeskiego, polskiego czy rosyjskiego. Co trzeba podkreślić, rozpoczęcie prac nad ich budową nie było równoznaczne z ich udostępnieniem, do którego dochodziło zwykle po upływie wielu lat.

Dla przykładu, choć prace nad Narodowym Korpusem Języka Czeskiego (czes. *Český národní korpus*) rozpoczęły się w 1994 roku, to on sam (pierwszy korpus referencyjny) został udostępniony na początku trzeciego tysiąclecia.

Narodowy Korpus Języka Polskiego to wspólne przedsięwzięcie Instytutu Podstaw Informatyki PAN, Instytutu Języka Polskiego PAN, Wydawnictwa Naukowego PWN oraz Zakładu Językoznawstwa Komputerowego i Korpusowego Uniwersytetu Łódzkiego. I choć sam korpus został udostępniony w 2012 roku w rezultacie realizacji projektu w latach 2007–2011, to wspomniane instytucje gromadziły dane korpusowe i opracowywały narzędzia korpusowe już w latach dziewięćdziesiątych.

Publiczne udostępnienie do Narodowego Korpusu Języka Rosyjskiego (ros. Национальный корпус русского языка) nastąpiło w 2004 roku, a poprzedziły go prace prowadzone od 2001 roku przez wiele instytucji, których działania koordynował Instytut Języka Rosyjskiego im. W.W. Winogradowa Rosyjskiej Akademii Nauk w Moskwie.

Mimo że przywołałem tylko te trzy słowiańskie korpusy narodowe, nie zapominam o innych, których najpełniejszy chyba opis znajdziemy w monografii *Praktyczny przewodnik po korpusach języków słowiańskiej* (2014) pod redakcją Mileny Hebal-Jezińskiej. Opracowanie to unaocznia, jak zróżnicowany jest zarówno stan badań korpusowych, jak i samych korpusów poszczególnych języków słowiańskich.

Uwzględniając okresy, w jakich kształtowały się inne dziedziny, na przykład pragmalingwistyka, można mówić o początkowym stadium rozwoju lingwistyki korpusowej. Jednak w odróżnieniu od poprzednich metodologii, które czerpały z osiągnięć innych dyscyplin (na przykład psycholingwistyka z psychologii, socjolingwistyka z socjologii), wnosi ona do prowadzonych analiz coś zupełnie nowego — metody ilościowe i empiryczne dotąd rzadko stosowane w badaniach lingwistycznych, a tak popularne w naukach ścisłych. W ten sposób lingwistyka korpusowa stanowi nowe podejście w metodologii badań — podejście statystyczne. Ponadto korzystanie z zasobów i narzędzi korpusowych umożliwia łączenie metod ilościowych i jakościowych.

Potwierdzenie wartości badań korpusowych znajdują Czytelnicy tego „korpusowego” numeru „Przeglądu Rusycystycznego” w poszczególnych jego artykułach, które obrazują prawdziwy wachlarz możliwości, jakie korpusy stwarzają badaczom języka. Wśród autorów znaleźli się językoznawcy zarówno tworzący korpusy językowe, jak i wykorzystujący je do celów badawczych. Zespala te wszystkie teksty obecność w każdym z nich języka rosyjskiego. Nie ma w tym przypadku, ponieważ jest to zgodne z profilem czasopisma adres-

## WSTĘP

wanego głównie do rusycystów. To właśnie z myślą o nich zostali zaproszeni do jego współtworzenia Autorzy, którzy zechcieli podzielić się swoimi doświadczeniami z pracy z zasobami i narzędziami korpusowymi. Wszystkim im — z Lublina, Olsztyna, Opola, Szczecina, Tampere, Warszawy i Katowic — chciałbym gorąco podziękować za współpracę i łączące nas przekonanie, że środowisko rusycystyczne takiego opracowania potrzebuje. Mam tu na uwadze zwłaszcza młode pokolenie rusycystek i rusycystów — badaczy, którzy poszukują jeszcze swojej tożsamości naukowej i którzy może znajdą w korpusach nie tyle obiekt naukowej fascynacji, ile skuteczną metodę badania języka stanowiącą rozwiązanie alternatywne wobec tradycyjnych modeli teoretycznych i koncepcji metodologicznych.