

Wykorzystanie korpusu tekstów w badaniach słownictwa i słowotwórstwa gwarowego (na przykładzie polszczyzny brasławskiej)

Celem niniejszego opracowania jest przedstawienie różnych możliwości wykorzystania korpusu tekstów w badaniach słownictwa i słowotwórstwa polszczyzny gwarowej na Brasławszczyźnie. Gwary polskie na badanym obszarze są dziś bardzo rozproszone, co sprzyja silnym interferencjom białoruskim i rosyjskim. Najstarsze i średnie pokolenie wciąż zachowują swój język, chociaż najmłodsze pokolenie rzadko używa języka polskiego. Na Brasławszczyźnie zachowały się liczne rodziny, a także nieliczne wsie, w których językiem prymarnym jest język polski w jego regionalnej odmianie. Oprócz tego ludność w codziennych kontaktach posługuje się językiem rosyjskim lub gwarą białoruską oddziałującymi na polszczyznę. Gwary więc pozostają środkiem nawiązywania kontaktu między pokoleniami. Korpus pozwala w szczególności na stwierdzenie frekwencji użycia wyrazu w całym zbiorze i u poszczególnych informatorów (Rieger 2014: 40).

Podstawę materiałową opracowania stanowią teksty gwarowe, których autorzy reprezentują najstarsze pokolenie, żyjące na początku XXI wieku. Większość tekstów pochodzi z korpusu elektronicznego, który sporządziłam za pomocą oprogramowania AntConc. W artykule przedstawię próbę analizy współwystępujących wyrazów rodzimych i zapożyczonych w polszczyźnie brasławskiej na przykładzie pary synonimicznej *budować ~ stroić* oraz ich derywatów. Pokażę także możliwość stosowania korpusu w badaniach frekwencji dubletów słowotwórczych oraz formantów. Czasownik *stroić* (por. ros. *cmpлоумь*) i jego derywaty można traktować jako efekt wpływów rosyjskich, w znacznej mierze nowszych (sowieckich), nawet jeśli został on zanotowany w języku starszego pokolenia. Wyraz ten i jego derywaty są też znane w gwarach polskich. Przypuszczam, że preferencje co do wyboru konkurujących między sobą wyrazów w parze synonimicznej (wyraz ogólnopolski ~ wyraz obcy), jak i częstotliwość użycia wyrazu zależy w pewnym stopniu od grupy wiekowej, co wymaga odrębnego zbadania w grupach pokoleniowych.

Na pary synonimiczne typu wyraz rodzimy i wyraz zapożyczony zwróciła uwagę N. Ananiewa (Ананьева 1990). Zestawianie zapożyczeń i synonimicznych wyrazów rodzimych (także archaizmów i wyrazów ogólnopolskich) stosowano w szeregu słowników i monografii¹. Natomiast zestawienie statystyczne zapożyczeń i wyrazów rodzimych w gwarze Maćkowiec na Podolu przedstawiła A. Krawczyk-Wieczorek za pomocą elektronicznego korpusu tekstów gwarowych w programie Poliqarp (Krawczyk 2007, 2011). Jednym z celów Autorki było przedstawienie proporcji między udziałem poszczególnych warstw słownictwa oraz wybranych jednostek leksykalnych we wszystkich tekstach oraz u poszczególnych informatorów. Obliczenia pokazały, że zapożyczeń w rejestrze jest bardzo mało, a w tekstach – jeszcze mniej. Korpus tekstów, sporządzony przez A. Krawczyk-Wieczorek również w systemie Poliqarp, wykorzystał J. Rieger w monografii *Słownictwo polszczyzny gwarowej na Brasławszczyźnie* w celu pokazania częstotliwości występowania wybranych wyrazów w parach wyrazowych, a także funkcjonowania zapożyczeń w gwarze. Jak wskazuje Autor, korpus tekstów pomaga przy zestawianiu słownika, pozwala przywołać konteksty użycia wyrazu i je rozróżnić, pomaga także w szybkim odnalezieniu synonimów, choć trzeba wtedy analizować każde użycie (Rieger 2014: 40–46). Poniżej podaję wybrane przykłady opracowania frekwencji wyrazów w parach:

białorutenizmy ~ wyrazy rodzime

hulać/pohulać 21 ~ *tańczyć/potańczyć* 94. Znaczna przewaga form polskich związana jest zapewne z brus. gw. *танцава`аць*.

rusycyzmy ~ wyrazy rodzime

daże 16 ~ *nawet* 158. Można mówić o znacznej przewadze wyrazu rodzimego, ale trudno nie zauważyć, że użycie rusycyzmu wyraźnie się zaznacza.

regionalizmy/zapożyczenia białoruskie lub rosyjskie ~ wyrazy rodzime

nadziewać i *nadziac* 9 ~ *ubierać* i *ubrać* 75.

Inne

sioło 3 ~ *wieś* oraz *wioska* i *wiosieczka* (razem) 322. Stare zapożyczenie, a później dialektyzm kresowy *sioło* (z akcentem na przedostatniej zgłosce od końca) jest prawie niezauważalny przy wyrazie *wieś* i derywatach. Forma *wioska* ma oparcie w tak samo brzmiącym wyrazie gwarowym białoruskim *вёска*.

¹ W słowniku Oleszkowiec I. Cechosz-Felczyk (2004), w słowniku Satanowa R. Modzelewskiej-Rewuckiej i J.Riegera (2008), w słowniku J.Riegera, I. Masojć i K. Rutkowskiej z Litwy (2006), w monografii O. Zakhutskiej o słownictwie wsi Siaberka na Wołyniu (2015).

W swoim opracowaniu posługuję się programem korpusowym AntConc² autorstwa Lawrence Antony'ego, który jest publicznie dostępny, szybki, łatwy w użyciu (w porównaniu z PoliQarpem, wymagającym znajomości instrukcji CQL³). Program ten ma wiele możliwości (m.in. pokazanie konkordancji, kolokacji, listy frekwencyjnej, słów kluczowych, dystrybucji) i może być wykorzystany jako efektywne narzędzie do badań słownictwa i słowotwórstwa. Na korpus wykorzystany w artykule składają się 64 teksty (ok. 70 godzin rozpisanych nagrań). Warto przy tym zwrócić uwagę, że precyzyjność wyszukiwania zależy od ujednolicenia zapisu półfonetycznego we wszystkich tekstach i usunięcia zbędnych symboli.

W celu pokazania możliwości wykorzystania korpusu w badaniach wybrałam parę synonimiczną *budować* ~ *stroić* oraz ich derywaty. Frekwencja czasowników *budować* oraz *stroić* wraz z derywatami w tekstach jest prawie jednakowa.

1. Konkordancja wyrazów *budować* (51) ~ *stroić* (40) oraz ich derywatów

W polszczyźnie brasławskiej wyraz *stroić* występuje w różnych znaczeniach: 1. 'przygotowywać, urządzać, szykować' (np. *ubranie, choinkę, jedzenie*) 2. 'organizować, robić' (np. *kołchoz, światło*) 3. 'budować' (np. *dom, chatę*). Niekiedy musimy zapoznać się z szerszym kontekstem, żeby podać dokładną frekwencję wyrazu w żądanym znaczeniu.

Za pomocą korpusu wynotowałam następujące derywaty od wyrazu *budować*: *nabudować* (2), *niedobudować* (2), *odbudować* (10), *pobudować* (24), *przybudować* (1), *zbudować* (10), *wybudować* (10), pojawiają się także formacje rzeczownikowe *zabudowa* (2) i *zabudowania* (7). Odpowiednio wynotowałam derywaty od wyrazu *stroić*, które pojawiły się w tekstach: *nastroić* (7), *postroić* (49), *przystroić* (4), *przestroić* (1), *rozstroić* (1), *obstroić* (2), *odstroić* (5), *wystroić* (3), *zastroić* (1), także formacje rzeczownikowe *postrojek* (por. ros. *ноcmпlоўкa*). Nie uwzględniam jednak derywatów *pierestrojka* oraz *ustrajać/ustroić* 'organizować, urządzać' występujących w tekstach w innym znaczeniu.

Sprawdzając wystąpienie par synonimicznych w korpusie, można zobaczyć, jakie formy są częstsze: *pobudować* (24) ~ *postroić* (49), *nabudować* (2) ~ *nastroić* (7), *przybudować* (1) ~ *przystroić* (4) – z przewagą wyrazów zapożyczonych, *odbu-*

² AntConc 3.4.4w (Windows) Developed by Laurence Anthony Faculty of Science and Engineering Waseda University, Japan, dostęp: <http://www.antlab.sci.waseda.ac.jp/software.html>

³ Corpus Query Language pozwala na wyszukiwanie elementu (słowa), por. zapytanie [base=*.budynek:* & i=iS & k=kN & recyt!=recytT & npol!=npolT & nol!=nolT & npol0!=npol0T & cyt!=cytW & cytat!=cytatT] pozwala na wyszukiwanie wszystkich wystąpień nie w komentarzu, nie w tekście recytowanym, nie w tekście niepolskim, nie w cytacie.

dować (10) ~ odstroić (5), wybudować (10) ~ wystroić (3) – z przewagą wyrazów rodzimych. Trzeba też wziąć pod uwagę formy z tzw. akaniem (np. *pubudować* ~ *pastroić*), gdyż jest ono charakterystyczne dla rosyjskiego i białoruskiego języka literackiego, dla znacznej części gwar białoruskich i rosyjskich, i także dla gwar polskich na Braślowszczyźnie.

Wyszukiwanie wyrazów w korpusie zilustruję na przykładzie pary synonimicznej *POBUDOWAĆ* ~ *POSTROIĆ*. W podanej frekwencji uwzględniłam wystąpienie danego wyrazu tylko w informatora, nie biorę natomiast pod uwagę użycia go przez eksploratora.

POBUDOWAĆ (21) + PABUDOWAĆ (3)

The screenshot shows the AntConc 3.4.4w (Windows) 2014 interface. The search term is 'pobudowac'. The results are displayed in a table with columns for Hit, KWIC, and File. The search is sorted by Level 1 IR.

Hit	KWIC	File
8	na zamka zylil, i , byl zamek sam , pobudowany , i du nas [Kszyzaki] oni tam , przys	709 (710A) GirD
9	, i pierszy chram , do tu , jaki byl pobudowany (on byl katolicki chram , na , tesz na	709 (710A) GirD
10	, nu dia zamag do tego czasu tam pobudowany , (on wot , stal nuscyc sie za to	709 (710A) GirD
11	, on nie taki stary , no alle , nie pobudowany posile wojny , za to ze jag [GES] ,	709 (710A) GirD
12	ji wszystko , i na tym miejscu ojciec pobudowal te kapliczki , i długi czas mówili że ,	709 (710A) GirD
13	, pliotki tam , no , tam dzie kościół teras pobudowany na tym , tam jakaż była jak knyznica	709 (710A) GirD
14	i taka cerka , no on'a , cerkof pobudowana w takim , bizantyskim stylu jak to na	709 (710A) GirD
15	, i , my tam stali na , tam koszary pobudowane jest , tam nocowali my , i tam byli ,	713B GawS17m.
16	co to za budynki tutaj stoja takie pobudowane? Rozwalone juz.] WL: a a nut a	716 GawWL36m
17	znalezli tyko fundamenty później pobudowali/ni tam n'a_wsi y	717 Gaw Niur.3t
18	i tam i magazyn , magazynsz pobudowany później , później co jeszcze tam aj cho	744B (745) Ccl5
19	, tutaj [...] gdzie dom tam. To ten , co pobudowany, to teraz, a pałac stoi, to on	749.txt
20	Guci, byli takie wielka ziemielne, szkola była pobudowana, duża była, piękna szkola była, a teraz	750 KoBT2k.txt
21	juz, teraz juz nie ma tego, w [...] pobudowali kościół, choć jeszcze nieskonczony, ma	757A_1 AchA2k
22	... aj, za dwa lata, za dwa lata, pobudowali, nie pamientam juz. Ale dwa lata juz, 757A_1 AchA2k	

Search Term: pobudowac
 Concordance Hits: 3
 Total No.: 64
 Files Processed: 6

POSTROIĆ (33) + PASTROIĆ (16)

The screenshot shows the AntConc 3.4.4w (Windows) 2014 interface. The search term is 'postroic'. The results are displayed in a table with columns for Hit, KWIC, and File. The search is sorted by Level 1 IR.

Hit	KWIC	File
16	szkola , jak dom tak szkola , później juz postroili , moze u pięćdziesiąt uamym roku , tys	600A (598A)
17	, as pallia sia laznia. A cieniżka laty postroili laznia. Wzieli, u taka na guze klatki,	600B Dwe22
18	: a fedyj ja zawożła wid'k , i postroili [niewyraznie...] bez batarej postroji	603B.txt
19	'k , i postroili [niewyraznie...] bez batarej postroili , nu ji tak ladnie gada po pol'	705 KW1Wm
20	białorusku tutaj/ni tej teras me postroico balatanka juz 'yadnożyli [i] adnacz	705 KW1Wm
21	ten dom był juz tedy postroiony , bez wiele lat przezdo nu	705 KW1Wm
22	ja i derala i ty chate ta postroili eszaz je jao mójce jak tam spalił	705 KW1Wm
23	i wszedzie wiazło jeszcze bedzi naz'ekac postroili chate ja tako jągiedie i pinie ndzy	705 KW1Wm
24	zrobili i i wszystko tam co jakiego postroili zyjim w swojej nazyl potem	707 (068) G
25	robijo , juz kolkta jedne , z drzewa , tak postroio struzkami [i] , i kolasa ta krensi sie	707 (068) G
26	można kieny tyka , i w Dwjw'atach kościół postroiony nie tak dawno , przed wojno , ja czy	711 (710) Z
27	, wynajmiaz, gospodar, który dom stroje, a tak, postroji j'oni naz'lowe ludzi, domi tyko	716 GawW13
28	, edycki aly f'e by... o teraz postroili nu aj'ej, tyk, nastrojne teraz kości	716 GawW13
29	mieli gdzie staw'ot' nie bylo tych postroijek tedyz tu po... po'astawili ot cieie	

Search Term: postroic
 Concordance Hits: 15
 Total No.: 64
 Files Processed: 6

Możemy zaobserwować, że w gwarach polskich na badanym obszarze występuje zarówno wyraz rodzimy, jak również zapożyczony, przy czym często ten sam informator naprzemiennie używa form polskich i wschodniosłowiańskich (Teksty 359, 360, 716, 717, 749).

2. Rozkład konkordancji

Na przykładzie par synonimicznych sprawdź, czy poszukiwane wyrazy występują w każdym tekście i jak często pojawiają się w poszczególnych tekstach.

ODBUDOWAĆ (10) ~ Odstroić (5)

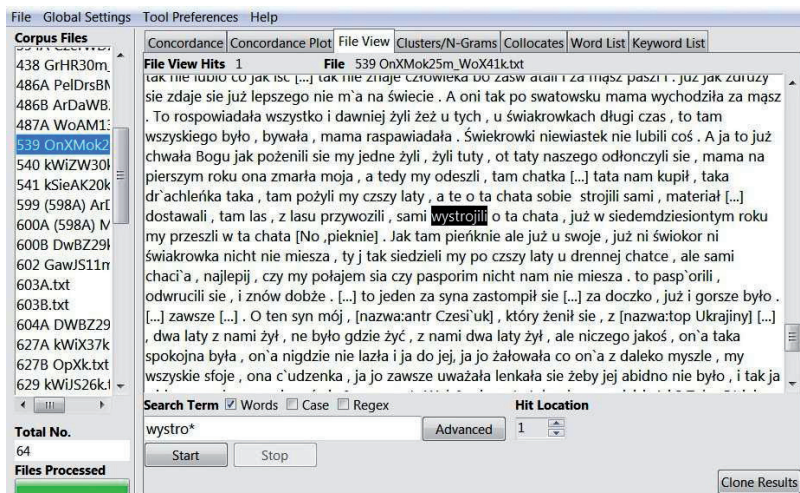
Wyraz *ODBUDOWAĆ* występuje w 8 tekstach, liczba użyć różni się minimalnie od 1, maksymalnie do 3 użyć. Natomiast *ODSTROIĆ* występuje tylko w 2 tekstach (u informatorów urodzonych w okolicach Dryświat kDrśMR22k, KoBT22k).

NABUDOWAĆ (2) ~ NASTROIĆ / NASTRAJAĆ (6)

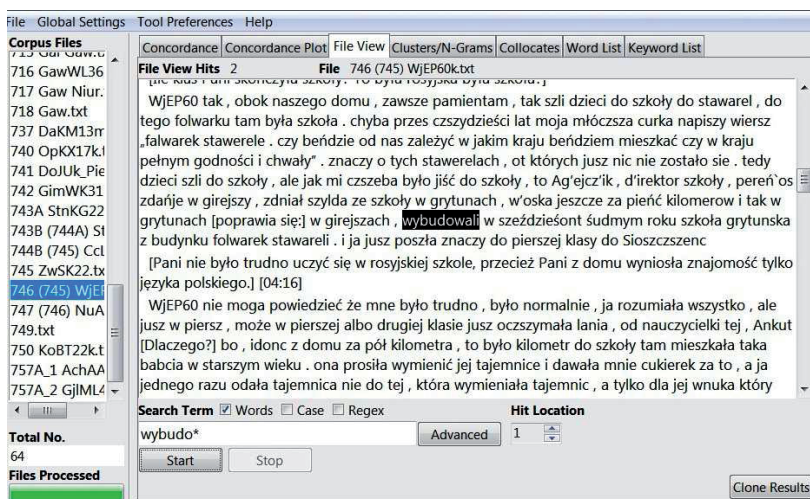
Wyraz *NABUDOWAĆ* występuje w 1 tekście (u informatorce ZaHA21k), a wyraz *NASTROIĆ* w 5 tekstach (3 razy używany przez informatorkę kWi-ZW30k, natomiast 2 razy w innym znaczeniu, por.: *Tkała, tkać będą jeszcze przyszła tam sąsiadka z pokryciem jednym i uot nastroiła jej te już tam no pracyndały te wszystkie; On mi rower nastroił, ja na roweru pojechała, przyjechała do chaty, użyć tych nie uwzględniam, podając frekwencję).*

3. Pełny tekst

Dzięki korpusowi można wyszukiwać wyrazy występujące u poszczególnych informatorów, np. *WYSTROIĆ* używa informatorka z Wozgielaniec na pn. zach. od Widz (WoX41k): „[...] dostawali, tam las, z lasu przywozili, sami wystrojili o ta chata, już w siedemdziesiątym roku my przeszli w ta chata” oraz informatorka z Widz (WiFD22k): „*tosz byli popalone wsi i wsio to wszystkie u Widzy zlicieli sie domy*” *wystrojili i żyjo [...]. Dom stroji dlia syna^e swego, dom trzeba wystrojć, zamysłół trzebaby poki jeszcze ja młody*”.



WoX41k w porównaniu z drugą informatorką jest młodsza, dlatego oddziaływanie języka rosyjskiego wydają się oczywiste (tu brak szczegółowej informacji o respondentce). Natomiast WiFD22k skończyła cztery klasy polskiej szkoły w Widzach i miała męża Polaka, z którym rozmawiała po polsku, ale miała też i dwóch synów, z którymi rozmawiała „po rusku” – tu interferencja wydaje się zauważalna ze względu na bliski kontakt z językiem rosyjskim.



Czasownik **WYBUDOWAĆ** występuje u ośmiu informatorów (DaKM13m, ZaHA21k, CzerWD27k, CcLS27k, kOpAB30k, WiWB31k, WjEP60k, DrjXk) zarówno najstarszych, jak i młodszych. Przykładowo informatorka WjEP60k używa w swoich wypowiedziach wyrazów rodzimych, po pierwsze dlatego, że jest nauczycielką języka rosyjskiego i polskiego, po drugie dlatego, że jej dzieci znają polski, jedna córka skończyła nawet studia w Polsce. A więc jak pokazują powyższe przykłady, kontakt z młodszym pokoleniem może z jednej strony podtrzymywać zachowanie form rodzimych, a z drugiej wpływa na szczerzenie się form zapożyczonych w polszczyźnie starszego pokolenia.

4. Klastry (zbitki słów z szukanym wyrazem)

Na przykładzie sporządzonej listy klastrów chcę sprawdzić, co właśnie budowali, czyli z jakimi rzeczownikami łączy się wyraz **budować** oraz jego derywaty (m.in. *budowa*, *budowanie*, *budownictwo*):

- *dom (domek) / mieszkanie / chlew / szopa / kuchnia / ściany*
- *kościół (kościółek) / kapliczka / chram*
- *szkoła / magazyn / miasto / wioska / elektrownia (elektrostanacja) / zamek / droga*

Wybrane przykłady użycia klastrów w trzech grupach:

- 1) *Na plibani tego modlili si, tedy tego post, zbudowali taka szopa, nu tak w tej szopie co trzy lat musi modlili sie DrśPWk20m. Wiezie Edmuś Kukj'an doski jak budował mieszkanie sobie w Szarkowszczyźnie, to*

na doski położył i rowerem, rowerem wozila sprzedawać kMiZK21k. Wybudowali we dwoje dom, na połowe tak rozdzielili i miszkają teraz tu CzerWD27k. O ten pokój był, i kuchnia byli odbudowane, a tamte wo, te te te pokoje byli naszymy jeszcze'lo nie dorobjona kGirWO40k.

- 2) To dawno dawno była, kościół tam zbudowany był, o to ja go pamientam; nabudowali kapliców i teraz, tam, ʘo dzie moja córka tam kaplica zrobjony ji pienkna ZaHA21k. I pierszy chram, ʘo tu, jaki był pobudowany ʘon był katolicki chram GirDrESz54k.
- 3) A potem już ruskie jak przyszli tak oni zaczęli budować tam i tam i magazyn, magazynsz pobudowany CcLS27k. Jak ja żyła jeszcze w Gimzach tam jeszcze w tamtym wengle to była szkoła ładana pabudowana, ja chodziła ale spaliła sie GimWK31k. Ale jak poszła u nas ta znaczy budownicwa tej elektrostancy, to wyrombali połowa lasu kWijF54k. Twój ojciec był przy budowie elektrowni GawJSz11m. W ostatnich dwu przykładach możemy zaobserwować różnice pokoleniowe w użyciu pary synonimicznej elektrownia ~ elektrosanca, najstarszy informator używa wyrazu rodzimego, a informator ze średniego pokolenia już używa wyrazu zapożyczonego.

Natomiast **stroić** (oraz derywaty) zwykle występuje w połączeniu z rzeczownikami:

- *chata / chałupy / dom (domik) / pokojczok / kwatery*
- *kościół / cerkwa / ołtarz*
- *brama / łaźnia / belka*

Wybrane przykłady użycia połączeń wyrazowych:

- 1) A pajechała zarabiać i tak mē dobrze wychodziła ta chata ta strojiła z tych zarabotków kWiwM10k. Tam, tam duzo było tych domow to też znaczyt idzie to teraz strojli sie drugie domy kWizW30k. Dwie chatki stojo i wso ty u w Gawranach też w magazyny też wo stojo [...] może any bywszy Aniela wyjechała u ten u Śnieszkus jak tu pastrojili tedy kwatere GawWL36m. Nō to o przestrojił takie, jēszcz'lo jedna taka pok'ojiczok taki, no no teraz sam pastrojil wszystko sobie, prawda ludzi pomagali mi MaMB37m. A te o ta chata sobie strojili sami, materiał [...] dostawali tam las, z lasu przywozili, sami wystrojili o ta chata, już w siedemdziesiontym roku my przeszli w ta chata WoX41k.
- 2) A da nas już obstrojili ten kościół słabodzki, ksiondz pomagał wszystko nam robić w baradzienickim kościele kMiZK21k. Przy Polsce strojony, w dwudziestym czwartym roku zaczeno strojić ten kościół; tu prawasławnych ta cerkwa odstrojona, ale tutaj z Opsa przyjeżdżajo, tam przyjeżdżajo, z Mieżyn przyjeżdżajo, a tut swoich może pieńć sześ ludzi tyko kDrśMR22k. Starym ludziom, to co to na rok, na rok za rok wszystko to

mieropryjacja wszystkie, to on tam oltasz stroi rok, tam zbierajo kDr-śMR22k.

- 3) *Belki straje^a... rombalyⁱ i stawili chata kWiWM10k. A u n'as była łaźnia, a łaźnia była, u n'as paliła sia łaźnia. A cienszkie laty postroić łaźnia DwBZ29k. Bramy jak nastrojo, stolik postawiony i teraz bramy robio i stolik postawio tutej kWiZW30k.*

Porównując połączenia wyrazowe z obu czasownikami, możemy zauważyć, że w wypowiedziach najczęściej wspomina się o budowie domu/mieszkania, kościoła/kapliczki i szkoły, a stroi się chaty/domy, kościół/cerkiew.

5. Kolokacje, N-gramy, słowa kluczowe

Na przykładzie wyrazu DOM uzyskałam 182 typy kolokacji (listę słów można sortować według frekwencji prawy/lewy rząd, statystyki, słów). Program też daje możliwość znalezienia w tekstach N-gramów (min. trigramy), czyli najczęstszych połączeń wyrazowych oraz kolokacji, pozwala sporządzić listę słów kluczowych oraz listę referencyjną dla nich.

Concordance	Concordance Plo	File View	Clusters/N-Gram	Collocates	Concordance	Concordance Plo	File View	Clusters/N-Gram	File View	Clusters/N-Gra			
Total No. of Collocate Types: 182		Total No. of Collocate		Total No. of Collocate Types: 182		Total No. of 182		Total No.					
Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate	Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate	Stat	Collocate
1	2	0	2	10.69364	zbudować	23	1	0	1	10.69364	agrzewa	8.69364	odbudować
2	1	0	1	10.69364	zarobili	24	2	0	2	10.10868	caglany	8.37171	zagroda
3	1	0	1	10.69364	zajęty	25	1	0	1	9.69364	wyglądał	8.37171	udź
4	1	0	1	10.69364	wyglądał	26	1	0	1	9.69364	wybudować	8.37171	stroje
5	1	0	1	10.69364	wielki	27	1	0	1	9.69364	wiekszy	8.37171	sprzedal
6	1	0	1	10.69364	tozpałac	28	1	0	1	9.69364	stroj	8.37171	sprzedali
7	1	0	1	10.69364	stodole	29	1	0	1	9.69364	rozwał	8.37171	postawjony
8	1	0	1	10.69364	stareńki	30	1	0	1	9.69364	przedać	8.10868	spat
9	2	0	2	10.69364	razbierali	31	1	0	1	9.69364	pastrojł	8.10868	rozwałł
10	1	0	1	10.69364	pokazeż	32	1	0	1	9.69364	opieki	8.10868	kupiony
11	1	0	1	10.69364	piworowicze	33	1	0	1	9.69364	ojcam	7.69364	zrobjony
12	1	0	1	10.69364	pastrojć	34	1	0	1	9.69364	ogzewa	7.69364	stroji
13	1	0	1	10.69364	ogarnęła	35	1	0	1	9.69364	mikulckiego	7.52372	mówiy
14	1	0	1	10.69364	odwieczny	36	1	0	1	9.69364	drachly	7.52372	bjaly
15	1	0	1	10.69364	mazalewskich	37	1	0	1	9.69364	borowskiego	7.37171	zapisany
16	1	0	1	10.69364	kusaczek	38	1	0	1	9.69364	atkupili	7.28765	duzy
17	1	0	1	10.69364	kupilab	39	1	0	1	9.69364	ante	7.23421	pusty
18	1	0	1	10.69364	kiedysie	40	2	0	2	9.37171	budował	6.99320	kupila

6. Frekwencja dubletów słotwórczych

Występowanie dubletów słotwórczych jest spowodowane specyfiką badanego obszaru, na którym współfunkcjonuje kilka języków, będących w ciągłym kontakcie. Określeniem *dublety słotwórcze* posługuję się za M. Sajewiczem: „są to nazwy posiadające to samo znaczenie słotwórcze, oparte na tych samych podstawach słotwórczych, lecz utworzone za pomocą różnych formantów” (Sajewicz 1989: 267). Na podstawie korpusu tekstów możemy opracować frekwencję dubletów słotwórczych (np. *dziecko* 82 ~ *dzieciak* 7 ~ *dziecionok/dziecionek* 24 ~ *dzieciątko* 6 ~ *dzieciuk* 3) z pokazaniem formantów charakterystycznych oraz często używanych w polszczyźnie braśławskiej czy dla poszczególnych idiolektów. Konkurencję form rodzimych i zapożyczonych na przykładzie nazw istot młodych omówiłam w swoim artykule⁴.

Możemy sprawdzić także frekwencję form deminutywnych w porównaniu z formami neutralnymi występującymi w gwarach, por. *żona* 142 ~ *żonka* 24, *pokój* 52 ~ *pokoik* 8 ~ *pokoiczek (pokoiczok)* 2, *lampa* 39 ~ *lampka* 2 ~ *lampeczka* 4, *krowa/krowy* 360 ~ *krówka/krówki* 32, *burak* 4 / *buraki* 40 ~ *buraczek (buraczok)* 3 / *buraczki* 10. Stwierdzenie o szerzeniu się form deminutywnych w wypowiedziach informatorów możemy potwierdzić na tle form neutralnych.

7. Wyszukiwanie formacji przez formant

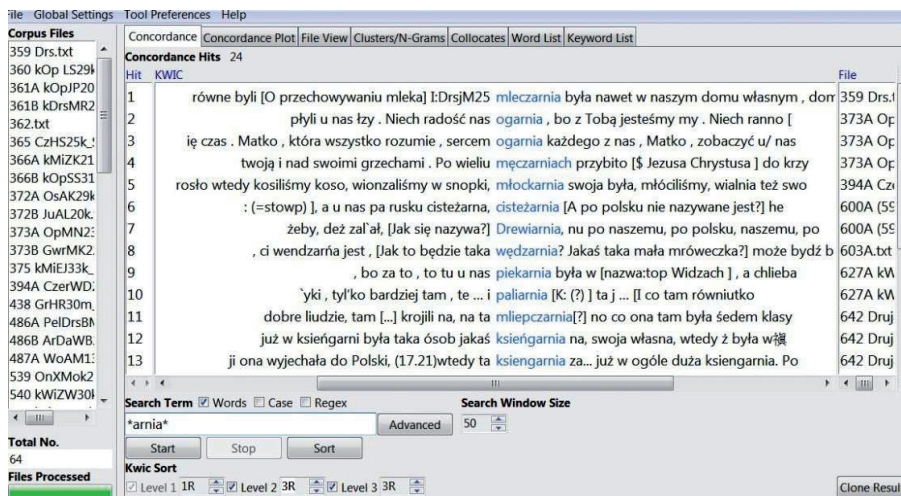
Korpus ułatwia poszukiwanie wyrazów z poszczególnym formantem. Na przykładzie sufiksu *-arnia* można spróbować wyszukać w tekstach wszystkie formacje (wśród 24 formacji rzeczownikowych znalazły się 3 czasowniki, np. *ogarnia*).

W języku ogólnopolskim formant *-arnia* należy do bardziej produktywnych w tworzeniu kategorii nazw miejsc (Grzegorzczukowa, Pużynina 1998: 413, 447). W badanym materiale formant *-arnia* tworzy formacje określające przeważnie zabudowania gospodarskie, ale też inne, por.

- nazwy miejsc/pomieszczeń

DREWIARNIA 1: *Jak drzewa składać żeby (...) drewnia, nu po naszymu, po polsku* MaMB37m

⁴ Referat pt. *Nazwy istot młodych w gwarach polskich na Braśławszczyźnie – konkurencja form rodzimych i zapożyczonych* został wygłoszony na Ogólnopolskiej Konferencji Naukowej „Małe języki Europy”, organizowanej przez Instytut Filologii Słowiańskiej UWr we Wrocławiu (23 maja 2016 r.). Artykuł został oddany do druku.



KSIEGARNIA 4 || KSIEŻARNIA 1: Tak już w księgarni była taka osob jakas księgarnia na swoja własna, wtedy ż była własność DrjXk. Książki wyda-wała z tego (...) on'a i teraz on'a w księżarni tej pracuje GrES32k.

MIASIARNIA 1: Na Płaszowie oni miaszkali, miał swoja miasarnia dzie wendlina wyrabiał, kiełbasy wyrabiał, i miał gospodarka jaka DaKM13m (< ros. м'ясо, brus. м'яса, ros. pot. мясарня)

MLECZARNIA 8: Osiemnaście krów mieli krów może, dużo mleka dawali, w Wanikiszkach była mleczarnia KoBT22k. Mleczarnia była nawet w naszym domu własnym DrsjM25k. Jak on na ladowce pracował w mleczarni DzJB27m. Zaraz wstała rano, padojiła tych krów, siostra do do mleczarni tam gdzie to taki, zawiozła jakim tam kolaskaj ji co ArDaWB29k. Tatus to moj nie poszet do kałchozu, on tak zrobił, że ůon wyszeł i pracował u Widzach w mleczarni kGir- WO40k. MOCZARNIA 1: Ona już niby to chowała sie w maczarnie^a alesz ludzie podpowiedzieli szyso jedno zabralyⁱ rosztrelaly GawTU28k.

PIEKARNIA 2: A tu piekarnia była swoja a potēm tu i młyn jusz ta^e bylyⁱ zniszczu^one wszystko WiAN20k. To tu u nas piekarnia była w Widzach, a chleba dostać nie można, to my z bratem sieńdzim na konia, a droga drewna była, przy-jedziem w nocy, kupim tego chleba kWiX37k.

PIWIARNIA 3: Zajeżdżajo tam do tego, ci tam piwjarnia była zajechali, i tak ůupodobał syn ta dziewczka ZaHA21k. Bo moje cjocia targowała piwiarnie miała i tam wódka a ja kupowała^e ze skljepa jej nosiła^e kWijsz19k.

SPİZARNIA 2: No tam chleb, siano skład`aliśmy, zboże. Jak sie wymłóci to do spizarni, syp`aliśmy CzerWD27k. Kiedyś byli na mieszkanie a teraz tak jak spizarnie zrobione DrjXk.

WĘDZIARNIA 1: *Kopjaki wybiera sie jakaś kość i tylic, samo mienso, ot obkłada sie marlio, wysusza sie a potem tym, łaźnia ci ci wendzarńa jest* GawJS23m

CISTEŻARNA? [STEŻARNIA] 2 ‘miejsce na stogi siana’: *Nu uot z siana, gdzie leży, tu nazywa sie, jak ad słowa stop, a u nas pa rusku cisteżarna, cisteżarnia* MaMB37m

W słowniku brasławskim J. Riegera notujemy inne formacje, reprezentujące ten typ słowotwórczy (na podstawie materiałów z kartoteki W. Werenicza): *owczarnia* ‘pomieszczenie dla owiec’ DrśW [SłBras, 185], *stożarnia* ‘bróg na siano lub słomę’ DrśW [SłBras, 251] (< *stóg*, por. brus. gw. *стажа́рня*, wyabstrahowany wtórnie z brus. zapożyczenia), *świniarnia* ‘pomieszczenie dla świń; chlew’ DrśW, WiW [SłBras, 263].

- nazwy sprzętów/narzędzi:

MŁOCKARNIA 1: *Nu sporo, trzeba było zasiać, żeby urosło wtedy kosiliśmy koso, wionzaliśmy w snopki, młockarnia swoja była, młóciliśmy* CzerWD27k.

Też *lichtarnia* ‘latarnia’ MM (por. brus. gw. *ліхтарня* ‘latarnia do prac w gospodarstwie’) [SłBras, 146].

- nazwy zbiorowe:

MĘCZARNIA 1 ‘męki, tortury’: *Po wieliu męczarniach przybito [Jezusa Chrystusa] do krzyża* OpMN23k. Też *kęściarnia* ‘zarośle, gąszcz’ GajA (< *kęst*, por. brus. brus. i ros. *кысм* ‘krzak’) [SłBras, 125].

- Inne:

PALIARNIA 1 ‘?’: *Tam na chliwie, tam mój posag. Kufry takie robisz, nu takie jak jaszczyki, tyl’ko bardziej tam, te (...) i paliarni* kWiX37k.

Łącznie w badanych tekstach pojawiło się 14 formacji, które występują 28 razy u 17 informatorów (4 mężczyzn, 13 kobiet). W polszczyźnie brasławskiej często występują dublety i synonimy słowotwórcze, jak np. *gospodarka* ~ *gospodarstwo*, *drewniarńik* ~ *drewniarńia*, *kuratńik* ~ *kurnik*, *kartoflaniszczce* ~ *bulbaniszczce*, a wykorzystanie korpusu może ułatwić wyszukiwanie tych nazw i sprawdzanie, czy występują one u jednego informatora, czy u kilku informatorów z jednej miejscowości oraz czy istnieje zróżnicowanie terytorialne formacji należących do poszczególnych kategorii.

* * *

Głównym zadaniem przedstawionego szkicu było ukazanie różnych możliwości wykorzystania korpusu w badaniach nad słownictwem i słowotwórstwem gwarowym. Po pierwsze, możemy poddać analizie pary synonimiczne (formy rodzime i zapożyczone) oraz dublety słowotwórcze, które często konkurują między sobą, a także pokazać, czy istnieją jakieś różnice w łączliwości poszczególnych par synonimicznych w polszczyźnie gwarowej. Pary wyrazów rodzimych

i zapożyczonych świadczą z jednej strony o zachowaniu tradycji języka polskiego, a z drugiej – o nowej fali zapożyczeń. Po drugie, możemy zobaczyć, jakie typy słowotwórcze są charakterystyczne dla systemu językowego, jakie przeważają, a jakie są sporadyczne. Po trzecie, możemy zbadać frekwencję i „moc” formantów (sufiksów i prefiksów) w formacjach rzeczownikowych i czasownikowych w poszczególnych kategoriach z pokazaniem formantów charakterystycznych oraz często używanych dla badanego obszaru lub też dla poszczególnych idiolektów. Wykorzystanie korpusu ułatwia pracę z dużymi zbiorami tekstów dzięki zautomatyzowaniu procesów poszukiwania. Elektroniczny korpus tekstów jako narzędzie badawcze stosowane do korpusowej analizy mowy pozwoli najpełniej opisać leksykę gwarową ludności wielojęzycznej i wielokulturowej.

Bibliografia

- Grzegorzczkova R., Puzynina J., 1998, *Rzeczownik* [w:] R. Grzegorzczkova, R. Laskowski, H. Wróbel (red.) *Gramatyka współczesnego języka polskiego. Morfologia*, Warszawa, s. 389–469.
- Krawczyk A., 2007, *Zapożyczenia leksykalne w sytuacji wielojęzyczności. Ukrainizmy i rusycyzmy w gwarze Maćkowiec na Podolu*, Warszawa.
- Krawczyk-Wieczorek A., 2011, *Słownictwo polskiej gwary kresowej na przykładzie Maćkowiec na Podolu. Charakterystyka funkcjonalna*, Warszawa [komputeropis].
- Rieger J., 2014, *Słownictwo polszczyzny gwarowej na Braślawszczyźnie*, Warszawa.
- Sajewicz M., 1989, *Dublety strukturalne w kategorii dewerbalnych nomina agentis w gwarach z pogranicza białorusko-krainskiego na Białostocczyźnie* [w:] S. Warchoł (red.), *Interferencje językowe na różnych obszarach Słowiańszczyzny*, Lublin, s. 265–271.
- Ананьева Н. Е., 1990, *Вариантность и синонимия в названиях частей тела в польских говорах Казахстана* [w:] J. Rieger, W. Werenicz (red.), *Studia nad polszczyzną kresową*, t. V, Wrocław, s. 185–204.

Use of Text Corpus in the Vocabulary and Word-formation Studies (in terms of Polish Dialects in Braslaw Region)

Summary

The paper presents various ways of using AntConc corpus programme as a tool for statistical analysis of vocabulary and derivatives in Polish dialects of Braslaw Region. The author makes an attempt to analyze the native and borrowed words on the example of the synonymic pair *budować* ~ *stroić* and their derivatives. And also shows whether there are any differences between certain synonymic pairs regarding collocations and word clusters. Synonymic word pairs indicate the preservation of the Polish language tradition, and on the other hand show a new wave of borrowings. Using text corpus

can help to detect what word-formation types are typical for Polish dialects system, which of them prevails and which are sporadic. And another thing, we can examine the productivity and frequency of noun and verb affixes in certain categories, showing characteristic and frequently used formants for Polish dialects in Braslaw Region or for individual idiolects. The use of electronic corpus facilitates the work with large sets of texts due to automatic search and helps to analyze the dialectal vocabulary of multilingual population.

Key words: text corpus, Polish dialects, vocabulary, derivation, loan words, synonymic pairs, Braslaw region

Słowa kluczowe: korpus tekstów, dialekty polskie, słownik, derywacja, zapożyczenia, pary synonimiczne, Braślowszczyzna