



**Vocabulary of the Polish dialects in  
Braslaw region in regard to frequency:  
a rivalry between the native and  
the borrowed forms**

**Yuliya Domitrak,  
University of Warsaw**

# The main goals:

to show some examples of text corpus as a database and a tool for language analysis in scientific research

to present the use of AntConc programming in the statistical analysis of native and borrowed words in the Polish dialects

# What is corpus?

## For what purpose can we use it?

I. It is a collection of texts we use for searching information, statistical analysis and hypothesis testing, checking occurrence or validating linguistic rules.

Modern corpora are:

- ✓ electronically stored and processed
- ✓ sustainable and representative
- ✓ structurally annotated
- ✓ searched with a help of special programs

# EXAMPLES OF CORPORA

**Language corpora** of Polish, Russian, Czech, English, Ukrainian etc.

- Narodowy Korpus Języka Polskiego – <http://nkjp.pl/>
- Национальный корпус русского языка – <http://ruscorpora.ru>
- Český Národní Korpus - <https://www.korpus.cz/>
- British National Corpus i Corpus of Contemporary American English – <https://www.english-corpora.org/bnc/> <https://www.english-corpora.org/coca/>
- Dabartinės lietuvių kalbos tekstynas – <http://tekstynas.vdu.lt/tekstynas/>
- Корпус української мови – <http://www.mova.info/corpus.aspx>  
Генеральний регіонально анотований корпус української мови (ГРАК) – [http://www.parasolcorpus.org/bonito/run.cgi/first\\_form](http://www.parasolcorpus.org/bonito/run.cgi/first_form)

**Parallel corpora:** Intercorp, ParaSOL, Polish-Russian Corpus UW

- Polsko-rosyjski i rosyjsko-polski korpus równoległy UW - <http://pol-ros.polon.uw.edu.pl/>
- Korpus Paralelny Języków Słowiańskich uniwersytetów Bernie i Ratyzbonie ParaSOL – <http://parasol.uni-be.ch>
- Korpus Paralelny Intercorp w ramach Czeskiego Korpusu Narodowego – <http://korpus.cz/intercorp/>



# Multimedial guides for Polish dialects in Poland and Lithuania

- Gwary polskie. Przewodnik multimedialny – <http://www.gwarypolskie.uw.edu.pl>
- Lietuvos lenkų tarmės - <http://www.lenkutarmes.flf.vu.lt/>



**Gwary polskie**  
Przewodnik multimedialny  
pod redakcją Haliny Karaś

### Mapa dialektów



**Menu wstępny**

- Start
- Wprowadzenie
- Podstawy dialektologii
- Opis dialektów**
- Dialekt masowiecki
- Dialekt wielkopolski
- Dialekt małopolski
- Dialekt śląski
- Nowe dialekty mieszane
- Leżakowian
- Autorzy
- Spis treści
- Litografia

Schematyczny podział dialektów polskich. Opr. A. Wiedonik na podstawie map: S. Urbańczyk, Żywa dialektologia polska, wyd. 3, Warszawa 1966, mapa nr 3 (Schematyczny podział dialektów polskich); Encyklopedia języka polskiego, pod red. S. Urbańczyka, Casselbaum, Wrocław - Warszawa - Kraków 1991, s. 413 (Mapa 1. Dialekty polskie (podział schematyczny)) z uzupełnieniami innych autorów.

© by Authors. Powered by Joomla!



**Lietuvos lenkų tarmės**  
2016 metų elektroninis leidinys, ISBN 978-809-459-803-8, Redaktorė Kristina Rutkoviška  
Vilniaus universitetas, Filologijos fakultetas, Polonistikos centras

Apie svetainę • Žemėlapis • Vietos • Naratyvai • Mokslo straipsniai • Žodynas • Bibliografija • Kronika



© Lietuvos lenkų tarmės, Rzecznictwo Polska Ministerstwo Spraw Zagranicznych, POMOC POLAKOM NA WSCHODZIE

Projekcija yra bendra nemieras © Lietuvos Respublikos Užsienio reikalų ministerija MŪJ, šioji projekto „Vilniaus ir Lietuvos Lenkai są gretos“ 2016 metų „gyvendinti“ fondų „Pomoc Polakom na Wschodzie“ lėpynas.

## O projekcie NKJP

### O PROJEKCIE NKJP

ZESPÓŁ

PUBLIKACJE

SŁOWA DNIA

TEKSTY KORPUSU

ZASTOSOWANIA

PODZIĘKOWANIA

NARZĘDZIA I ZASOBY

KONTAKT

Korpus językowy to zbiór tekstów, w którym szukamy typowych użyć słów i konstrukcji oraz innych informacji o ich znaczeniu i funkcji. Bez dostępu do korpusu nie da się dziś prowadzić badań językoznawczych, pisać słowników ani podręczników języków obcych, tworzyć wyszukiwarek uwzględniających polską odmianę, tłumaczy komputerowych ani innych programów zaawansowanej technologii językowej. Korpus jest niezbędny do pracy językoznawcom, ale korzystają zeń często także informatycy, historycy, bibliotekarze, badacze literatury i kultury oraz specjaliści z wielu innych dziedzin humanistycznych i informatycznych.

WYSZUKIWARKA KORPUSOWA  
IPI PAN

WYSZUKIWARKA KORPUSOWA  
PELCRA

**National Corpus of the Polish Language** (NKJP) is a database of 1.5 billion words or a balanced undercorpus - 300 million words. It reflects structure of reading in Poland (10% are spoken texts). There are 2 programs for searching information:

**Poliqarp** - a search engine with a fairly complex syntax of queries CQL, that allows you to find any sequences of words characterized by appropriate grammatical categories;

**Pelcra** - an easy-to-use program for searching the concordance of word forms and lexemes, as well as their collocations.

# CMOKIER 'flatterer, adulator'

Wyszukiwarka korpusowa PELCRA dla danych NKJP (Cytowanie)

cmokier\*\*

Maks. odstęp: 0    Zachowaj szyk:     Wyniki: 100    Czas    Profil    Excel    URL

Zaawansowane    Podkorpus: Cały

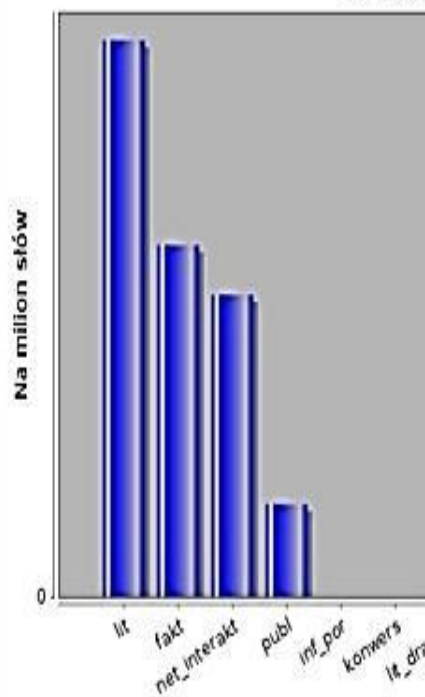
[Pomoc](#)

Przeszukiwany zbiór zawiera 1,524,696,745 słów. Znalaziono 37 akapitów pasujących do zapytania w 0.082s. Bieżąca strona zawiera 37 przykładów z 36 różnych tekstów.

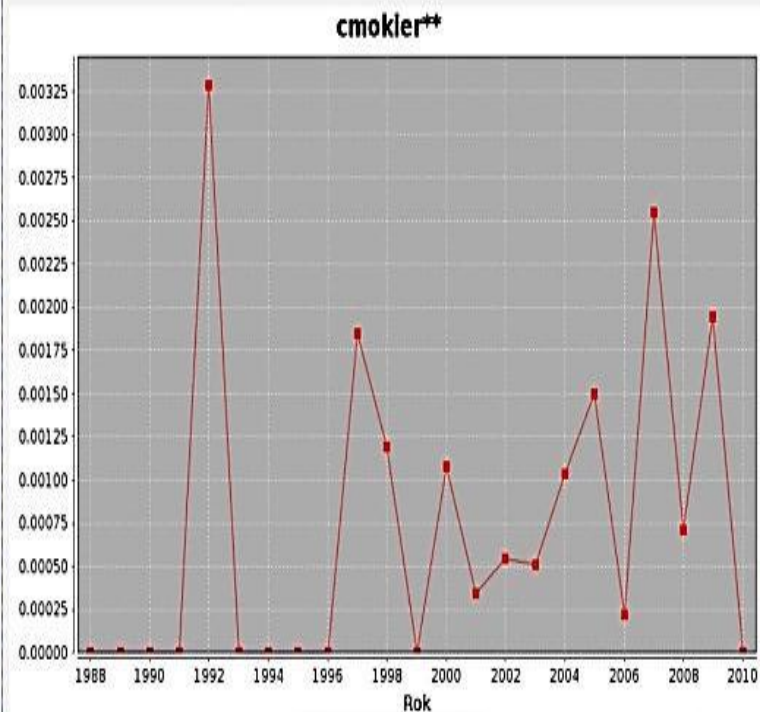
1.		cmokiera	a nie wyjecha	Usenet -- pl.sci.hi...
2.	Nie proszę Pana...nad niczym/nikim nie cmokam...nie jestem	cmokierem	Wszelako jeśli Pańska znajomość fizyki jest podobna jak muzyki	Forum Stanisława Mi...
3.	ochotą. [ Dodano: 09 Lis 2007 12:31 ] Pan sam może nawet	cmokierem	i nie jest, ale jak Pan widzi, że ktoś nie chce cmokać... To go	Forum Stanisława Mi...
4.	historii. Nie przeszkodziło to jednak skrajnym filosemickim	cmokierom	w robieniu w polskiej prasie reklamy nawet tak ohydnu	Usenet -- pl.soc.po...
5.	dwulicowy, dokonuje cudów, aby być obiektywnym. Myślę, że wbrew	cmokierom	RAZ-owym chwającym odwagę i ubolewającym, że się naraził, starsi	Forum Stanisława Mi...
6.	przychyłość. Toteż miał moc przyjaciół, wielbicieli i "	cmokierów	", jak mówili mniej życzliwi.	Gazeta Wyborcza
7.	Nie na wiele zdała się ta wiedza. Za popieranie "rewizjonistów i	cmokierów	" wyrzucono ją z PIW-u. Później też starano się spacyfikować ją,	Miałam dar zachwyty

## cmoker\*\*|cmokier\*\*

#	Typ	Liczba wystąpień (A)	Słowa w kategorii (B)	A/(B/1M)
1	net_interakt	22	456,520,815	0.048
2	publ	11	745,668,605	0.015
3	lit	3	33,754,248	0.089
4	fakt	1	17,807,497	0.056
5	inf_por	0	8,785,047	0
6	konwers	0	1,795,251	0
7	lit_dramat	0	207,669	0
8	lit_poezja	0	74,793	0
9	lit_proza	0	7,119,812	0
10	media	0	447,112	0
11	net_nicinterakt	0	119,924,305	0
12	nklas	0	2,914,447	0
13	qmow	0	92,546,016	0
14	urzed	0	7,161,072	0
15	nd	0	29,172,984	0
<b>Łącznie:</b>		<b>37</b>	<b>1,523,899,673</b>	<b>0.024</b>



#	Rok	wystąpienia	akapitów
1.	1988 0	10,534	0
2.	1989 0	16,109	0
3.	1990 0	12,291	0
4.	1991 0	25,794	0
5.	1992 1	304,144	0.003
6.	1993 0	329,534	0
7.	1994 0	417,518	0
8.	1995 0	419,001	0
9.	1996 0	448,266	0
10.	1997 1	541,852	0.002
11.	1998 1	835,922	0.001
12.	1999 0	1,371,446	0
13.	2000 2	1,852,251	0.001
14.	2001 1	2,881,174	0
15.	2002 2	3,675,398	0.001
16.	2003 2	3,938,533	0.001
17.	2004 3	2,899,653	0.001
18.	2005 7	4,660,002	0.002
19.	2006 1	4,483,486	0
20.	2007 6	2,354,746	0.003





# Corpus as a key tool for qualitative and quantitative analysis of natural language



II. Software helps to work with large collection of texts thanks to the automatic searching.

Corpus can be created from any set of your own resources

- ❑ according to discipline (literary texts or legal texts)
- ❑ according to particular author (works of Adam Mickiewicz or Bolesław Prus)
- ❑ according to certain époque (17th century Polish language, language of the young people)



# KORPUSOMAT

Created by scholars from IPI PAN: Witold Kieraś, Łukasz Kobyliński, Maciej Ogrodniczuk, Zbigniew Gawłowicz, Michał Wasiluk

<http://korpusomat.pl/>

KORPUSOMAT



KORPUSOMAT

Utwórz własny korpus językowy



MOJE KORPUSY  
JĘZYKOWE



UTWÓRZ KORPUS



PRZEŚLIJ OPINIĘ

O PROJEKCIE

It aggregates other existing tools: Morpheus 2 (analyzer and inflection generator for the Polish language); TermoPl (used to recognize terminology); Morphosyntactic Disambiguation (segmentation)

# ZAPYTANIA DO KORPUSU 'NOWY KORPUS'

Zapytanie  
[pos="pant"]

KONSTRUKTOR ZAPYTAŃ

METADANE ▾

STATYSTYKI ▾

Liczba wyników na stronę

10 ▾

Wyszukaj

Znaleziono 510 wyników.

28

Rezultat

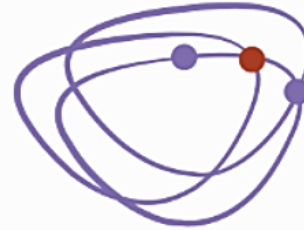
Prawy kontekst

Lp	Lewy kontekst	Rezultat	Prawy kontekst
1	do swego państwa; nakazał więc jak największe wysiłki i	zebrawszy [zebrać;pant;perf]	wojsko, wyruszył na czele jego 21 maja 1621 r
2	. Chodkiewicz, żołnierz wytrawny, który z najlepszej szkoły	wyszedłszy [wynieść;pant;perf]	. własnego geniuszu niepospolite złożył nieraz świadectwa, przedwcześnie z
3	Władysław przez Dniestr pod Chocim i przybyli Kozacy, ledwie	opędziwszy [opędzić;pant;perf]	się Turkom i Tatarom ( mirzy21 Kantymira22 i samego chana
4	[,dwa autorów przydłuższe zabawy tej marsowej sprawy wciąż	przeczytawszy [przeczytać;pant;perf]	* napisał je autor); Bartłomieja Zimorowica Pamiętka wojny
5	" (str. 51–142) sprawie chocimskiej	poświęciwszy [poświęcić;pant;perf]	. Panegiryzmem zwichnął nieco jej opis, bo tak coraz
6	śmy palili księgi owe Twardowskiego o ich ekspedycjach napisane32"	Nasarkawszy [nasarkać;pant;perf]	się na pobory, co kraj zubożyły, na Jana
7	nieraz się Potocki myli, czy z własnej winy,	pomieszawszy [pomieszać;pant;perf]	nazwy i rzeczy, czy z winy ogólnie powtarzanej bajedy
8	i najdrażliwsze dowcipy; — umarł też Arystyp najprozaiczniej,	zachorowawszy [zachorować;pant;perf]	podczas powrotu do ojczyzny. Za takimi to anegdotami upędał
9	np. który rano z Hussejnem, podskarbi trzecie miejsce	wziąwszy [wziąć;pant;perf]	. zaprasza itd. Form i słów narzeczowych używa dla

We can find lexical units by certain categories or grammatical classes (np. place names, adverbial participle); display a frequency list of results by a specific segment attribute or a graph showing the distribution of results due to metadata

# Common Language Resources and Technology Infrastructure

**CLARIN-PL**  
Common Language Resources and Technology Infrastructure



CENTRUM TECHNOLOGII  
JĘZYKOWYCH **CLARIN-PL**



## Czym jest CLARIN?

Celem CLARIN jest udostępnianie zasobów językowych oraz elektronicznych narzędzi do automatycznego przetwarzania języka naturalnego badaczom we wszystkich dyscyplinach naukowych, a w szczególności z dziedziny nauk humanistycznych i społecznych.

[Dowiedz się więcej ▶](#)



## Misja

Zasadniczy cel CLARIN to budowa narzędzi naukowych, które umożliwią rozszerzony dostęp do dużych zbiorów dokumentów tekstowych, nagrań języka mówionego i zasobów multimedialnych oraz ich zautomatyzowaną analizę.

[Dowiedz się więcej ▶](#)



## Struktura

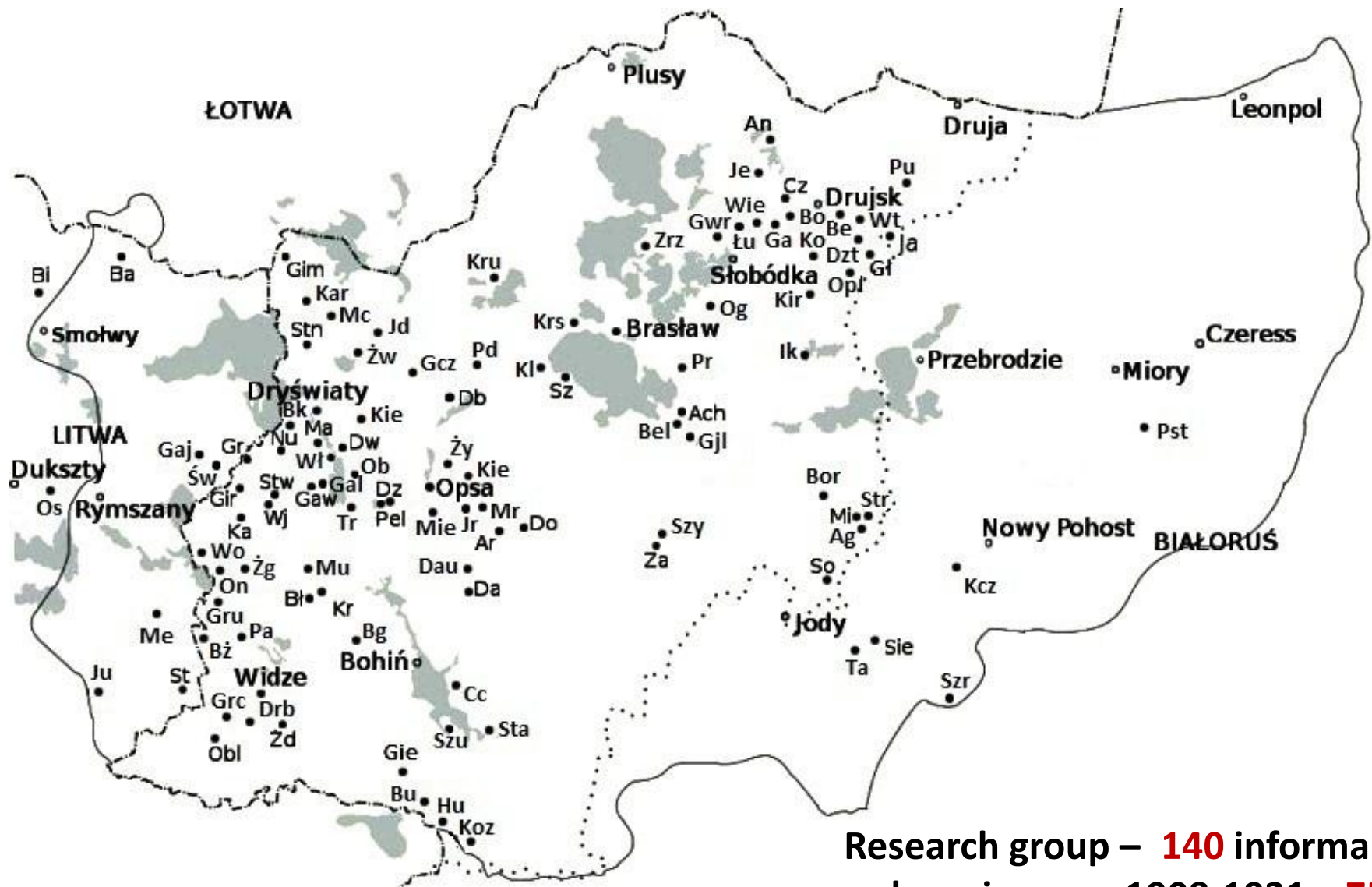
Jednostką zarządzającą CLARIN na poziomie europejskim jest konsorcjum państw o nazwie CLARIN ERIC, w którego skład wchodzi rząd każdego z państw oraz organizacje międzyrządowe.

[Dowiedz się więcej ▶](#)

It is a pan-European research infrastructure intended for the humanities and social sciences. Other programs for creating a corpus: Smyrna, Text Stat, ParaConc

<http://clarin-pl.eu/pl/uslugi/>

# BRASLAW REGION

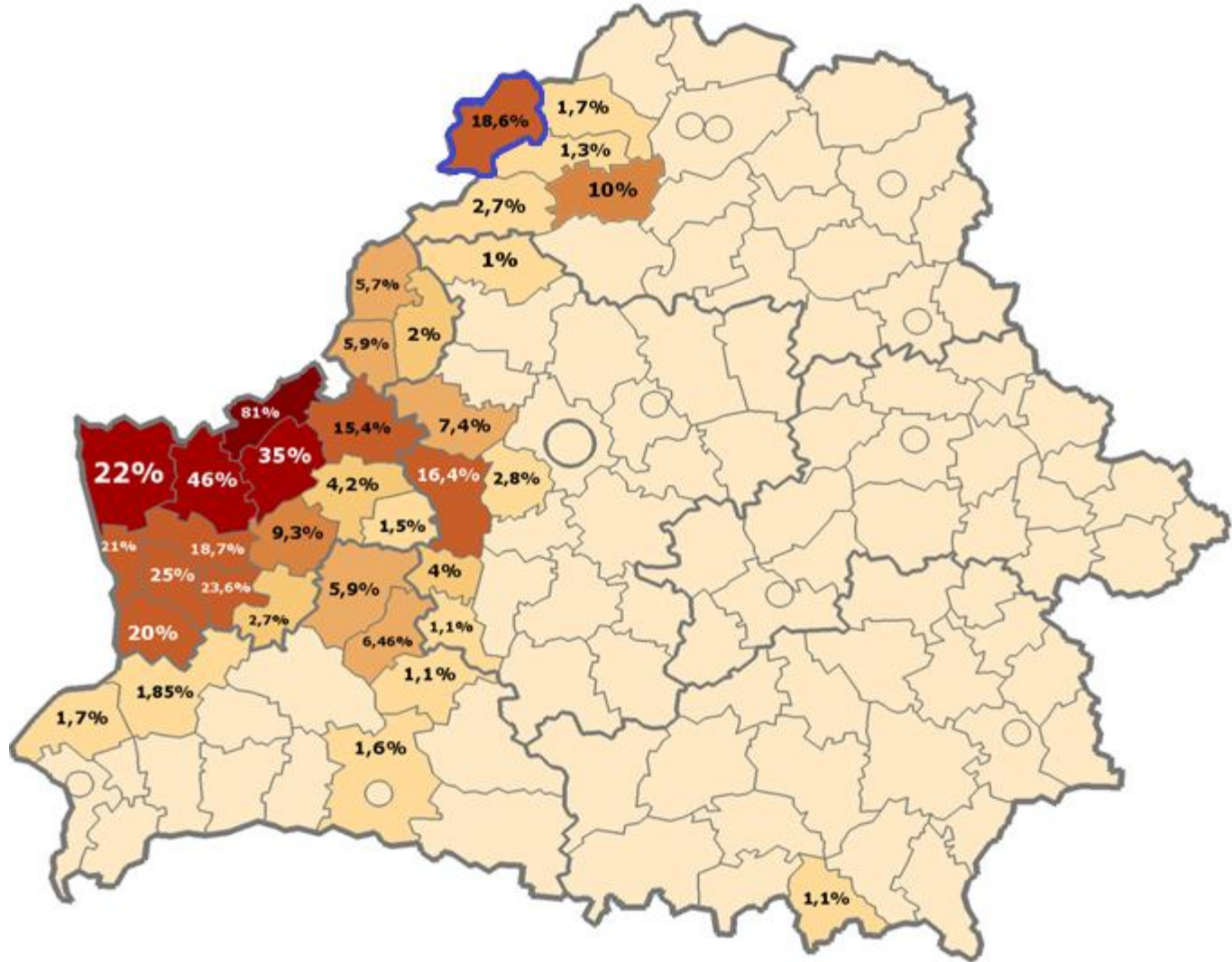


Research group – **140** informants

- born in years 1908-1931 – **75%**
- born in years 1932-1960 – **25%**



# Polish minority in Belarus (2009)

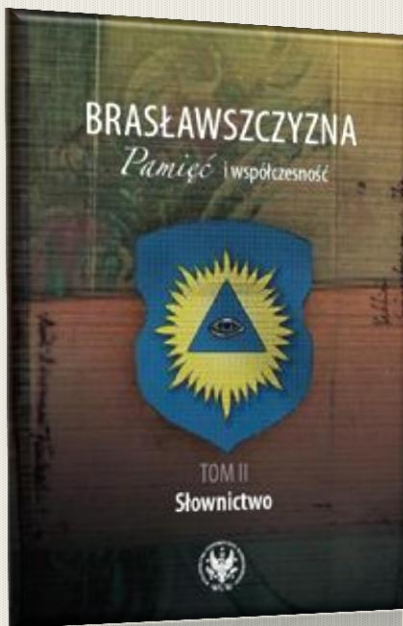


According to the population census from 2009 in Braslaw region live: **Belarusians – 64,6%, Poles – 18,6%, Russians – 14,2%**

## Research questions

- ❑ What benefits bring the results of research based on the electronic corpus in comparison with what has already been presented in scientific publications?
- ❑ Whether it is possible to accept some calculating and objective results above intuitive assumption and in particular show the changeability in the vocabulary of Polish dialects?





SMUŁKOWA E. (red.), *Brasławszczyzna. Pamięć i współczesność. Tom I Historia Regionu. Charakterystyka socjolingwistyczna. Świadectwo mieszkańców*, Warszawa 2011. Tom II. *Słownictwo dwujęzycznych mieszkańców rejonu*, Warszawa 2009.

History of the region, the sociolinguistic situation, testimony and memories of the population.

Dictionary of 'bilingual population of the region' and collection of spoken texts.

RIEGER J., *Słownictwo polszczyzny gwarowej na Brasławszczyźnie*, Warszawa 2014.

The language situation in the area, co-functioning of borrowed and native words (fragmentarily)

Vocabulary nearly 5000 words with texts sorted according to the subject catalogue.



# Corpus Program AntConc

Laurence Anthony's Website

Home Resume Publications Software Classes Photo Albums Links Contact

## Software



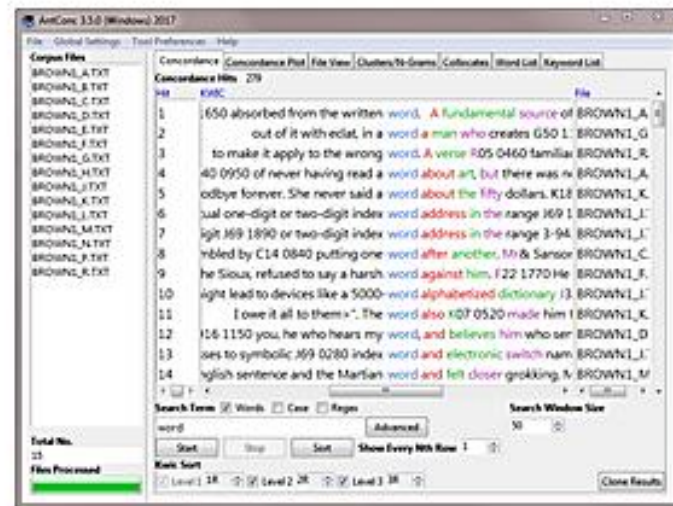
### AntConc

A freeware corpus analysis toolkit for concordancing and text analysis.

[[AntConc Homepage](#)] [[Screenshots](#)] [[Help](#)] [[License](#)]

#### Downloads:

- Windows (3.5.7)
- Windows 64-bit (3.5.7)
- Macintosh OS X 10.6-10.12 (3.5.7)
- Linux 32-bit (3.5.7)
- Linux 64-bit (3.5.7)
- Older versions



Laurence Anthony, Ph.D., Center for English Language Education  
in Science and Engineering, School of Science and Engineering,  
Waseda University, Tokyo, Japan

<http://www.laurenceanthony.net/software.html>



# **AntConc 3.4.4 – corpus program for text analysis that helps**



**to search a word in the whole corpus  
or in certain informants**



**to show concordance, frequency,  
distribution of words**



**to find clusters, collocations,  
N-grams, key words**

In my work I will try to examine the rivalry of native and borrowed words taking into account:

frequency of a word – its occasional or permanent, idiolect or common usage (e.g. ‘brick’ *cegły/cegiełki* 35 ~ *kirpicz* 2 GrES32)

parallel functioning of differential lexemes and their Polish equivalents in synonymous pairs and series at the individual and regional level (e.g. ‘uncle’ *wujek/wujaszek* 40, *stryj/stryjek* 15 ~ *dziadzia/dziadźka* 44, verbs with the derivatives ‘build’ *budować* 193 ~ *stroić* 153)

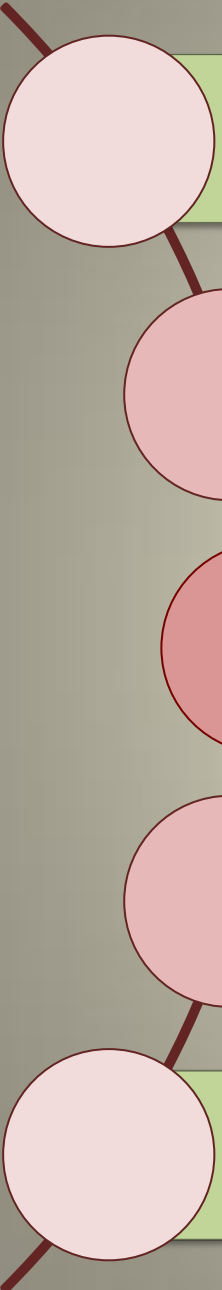
co-existence of word-formation doublets, showing the "power" of formants in noun and verbal formations (e.g. ‘child’ *dziecko* 91 ~ *dzieciak* 16 z 7 wsi ~ *dziecionok/dziecionek* 57 z 19 wsi ~ *dzieciątko* 5 ~ *dzieci<sup>l</sup>uk* 3 ~ *dziecina* 2)

# Problematic questions



- ✓ What kind of words may specify the character of a dialect or indicate the richness of foreign influence? What are reasons for that?
- ✓ What words should we take into consideration to determine the changeability of dialect?
- ✓ In which thematic groups the archaic and dialectal vocabulary is being kept longer; to which spheres do borrowings penetrate more often?
- ✓ What factors may influence on the functioning of vocabulary in certain idiolects?

# THEMATIC GROUPS OF VOCABULARY



**Family, blood relationship and kinship by marriage**

**House and domestic buildings**

**Food and cooking**

**Work in the field and on the farm**

**Man in a social environment**



## **The vocabulary that is not considered in the work:**

- ✓ vocabulary related to Soviet realities of that time, for example in the relations about collectivization, work in kolkhoz, names of professions (e.g. *brygadir, bugaltier, awtoławka, kamsamłot, kombajnior, pionierzy, sawchoz, sielsowiety, świnomatka, szczytawod*);
- ✓ medical vocabulary covering the names of diseases and methods of their treatment (e.g. *asma, gryża/hryża, dawlenie/dawlienije, opuchał, pozwanocznik, tuberkuloz, skarłatyna, ukoł*);
- ✓ the names of plants that are often used in home treatment (e.g. *czystacył, harlanki, durmanczyk, padaroznik, kalina, serdecznik*) or others names of mushrooms, flowers and trees, because these lexemes are usually borrowed;
- ✓ the vocabulary concerning religious life and Roman Catholic Church is native.

## FAMILY



- ‘family’ *rodzina* (505) in 76 places ~ *semja/siamja* (13) *siemiejstwo* (1) in 9 places ◇ brus. сям’я, ros. семья, сем|ейство
- ‘father’ *ojciec* (800), *tata/tato* (71) ~ *baćko/baťko* (11) in 8 places, *ociiec/otiec* (9) in 5 places, *papa* (47) in 24 places ◇ brus. gw. б|ацька, б|ацько, ros. па|па, от|ец
- ‘grandmother’ *babcia* (163) in 44 places, *babka* (26) ~ *bab|uszka* (21) in 8 places ◇ ros. ба|бушка
- ‘akin’ *rodzony* (58) in 30 places ~ *rodny* (9) in 7 places ◇ brus. ро|дны, ros. родн|ой

# Searching synonymic pair in corpus

## *rodzony ~rodny*

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

**Corpus Files**

AchTH23 -751  
AnGaLS27 -36  
ArDaWB29 -48  
ArRA33sz -752  
BaHC35 -L007  
BeXK26 -643.1  
BiKG19 -L003.  
BoJX25sz -739  
BrMDk40 -488  
CcLS27 -744,7  
CzHS25 -365.t  
CzWD27 -394.  
DbKM13sz -7.  
DoJU40sz -74  
DoXK31 -738.  
DrCX30pl -739  
DrjXM -643.txt

**Concordance Hits 65**

Hit	KWIC	File
24	oł proszki [...] ha?, jak, jak zaw`uć?, <b>rodzonyj</b> ?, w jedenastym. w Gawran	GawHS23,
25	nieniane, nie wiem ...nu, a? a ja tutaj <b>rodzona</b> , tutaj ja, ot i dożywam, trze	GawHS23,
26	ski Wiktor, też Wiktor. Bracia, bracia <b>rodzone</b> i obydwu podobna, ten sar	GawSP17
27	już też na pensji, usz jego mi sie <b>rodzony</b> brat na pensji o był ale ten	GawWL36
28	ki. A tut była jeszcze księdza siostra <b>rodzona</b> ona była tarciaarka była, też	GieWJ21 -6
29	był w domu i zaszed tam, jego bart <b>rodzony</b> mieszkał. I oni wtedy okrąż	GieWJ21 -6
30	ona taka nie żałuje piniędzy, jej brat <b>rodzony</b> był takim wielkim naczalnil	GieWJ21 -6
31	dzo ludzi. Mego wujka mamy brata <b>rodzoneo</b> było żona i pieć dzieci o	GieWJ21 -6
32	ot tutaj wo ona żyła. K2: Jejny brat	
33	była, dziadluk, a anja siostra jego,	
34	a ciocia Makowska, to mojej mamy	

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

**Corpus Files**

AchTH23 -751  
AnGaLS27 -36  
ArDaWB29 -48  
ArRA33sz -752  
BaHC35 -L007  
BeXK26 -643.1  
BiKG19 -L003.  
BoJX25sz -739  
BrMDk40 -488  
CcLS27 -744,7  
CzHS25 -365.t  
CzWD27 -394.  
DbKM13sz -7.  
DoJU40sz -74  
DoXK31 -738.  
DrCX30pl -739  
DrjXM -643.txt

**Search Term**  Words  Case  Regex

rodzon\*

Start Stop Sort

**Kwic Sort**

Level 1 1R  Level 2 2R  Level 3 3R

**Concordance Hits 10**

Hit	KWIC	File
1	ch braci i siostrów tam, i braty dwa <b>rodne</b> i tego jichne dzieci, tak tam d	ArRA33sz -
2	zie nie strzelajcie. Oni mówio, choć <b>rodnego</b> bačka zastrelim, i strzeliali i	DrXKk27, C
3	jak ojca. Antoni Antoni, n`u o Staški <b>rodny</b> brat umarszy ta umarszy. Dwi	GawCM52
4	t mego brata teściowa mańka i oń`i <b>rodna</b> siostry a mila młudsza hilary j	GawWL36
5	nienieli sie tak, jak chto mog i... swoj <b>rodny</b> język to już u n`as nieprawidk	GimWK31
6	sie Czesia taka. A wy nie tam nie <b>rodnili</b> sie, Stacha poumierala teraz,	GrES32 -7
7	ił sie. On z Repeńszczyzny dziadka <b>rodny</b> Kazik, no etat Kazik nie taki, a	KoEG24 -3
8	ajom dalej tam, jeśli kogo nie ma z <b>rodnych</b> , kiedy przyjadom, no to. A	KoEG24 -3
9	po rosyjsku? Ja rozmawiam swoim <b>rodnym</b> językiem, swoim rodnyim ję	kZaFJ27, Jc
10	im swoim rodnyim językiem, swoim <b>rodnym</b> językiem swoim własnym b	kZaFJ27, Jc

**Search Term**  Words  Case  Regex

rodn\*

Advanced 50

**Search Window Size**

# Names of kinship by marriage in Polish dialects of Braslaw region

Nazwy rodzime

Nazwy zapożyczone

## **brat ojca i matki**

<i>wujek, wujaszek</i>	<i>dziadzia, dziadźka</i>
<i>stryj, stryjek, stryjaszek</i>	

## **siostra ojca i matki**

<i>ciotka, ciocia</i>	<i>tiotia, tiotka</i>
-----------------------	-----------------------

## **rodzice męża i żony**

<i>teść, teściowa</i>	<i>świekrowa, świakrowka</i>
<i>swiakra,</i>	<i>świakr<sup>1</sup>ow, świekrucha</i>
<i>świokier, świokor</i>	<i>cioszcza</i>

## **brat i siostra męża**

<i>szwagier</i>	<i>dziewierz</i>
<i>szwagierka</i>	<i>załówka</i>

## **żona syna i brata**

<i>synowa, bratowa</i>	<i>niewiastka</i>
------------------------	-------------------

## **mąż córki**

<i>zięć</i>	<i>ziac</i>
-------------	-------------

## **dzieci siostry ojca i matki, brata matki**

<i>kuzyn, kuzynka</i>	<i>bracienik,</i>
<i>cioteczny brat/siostra</i>	<i>dwurodny/ trojurodny,</i>
<i>stryjeczny brat/siostra</i>	<i>dwoj<sup>1</sup>ur<sup>1</sup>odna/dwujerodna</i>
<i>siostrzeniec, bratanek</i>	<i>plemiennik/plemianik</i>
<i>siostrzenica</i>	<i>plemiennica/plemianica</i>

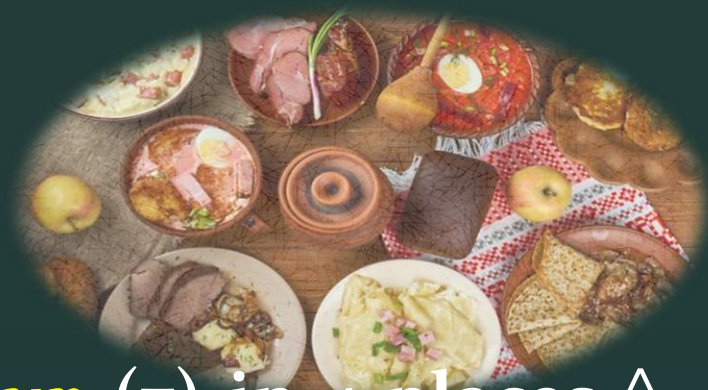




- ❑ On the basis of the prior analysis, we can indicate the geographical diversity of lexemes and formants.
- ❑ In the analyzed about 50 pairs and synonymic series, borrowings rarely dominate, sometimes the competitiveness of both systems is visible as a result of the languages interaction.
- ❑ Despite the predominant number of Polish words in most synonymic pairs, the presence of borrowings is clearly marked.



# FOOD AND COOKING



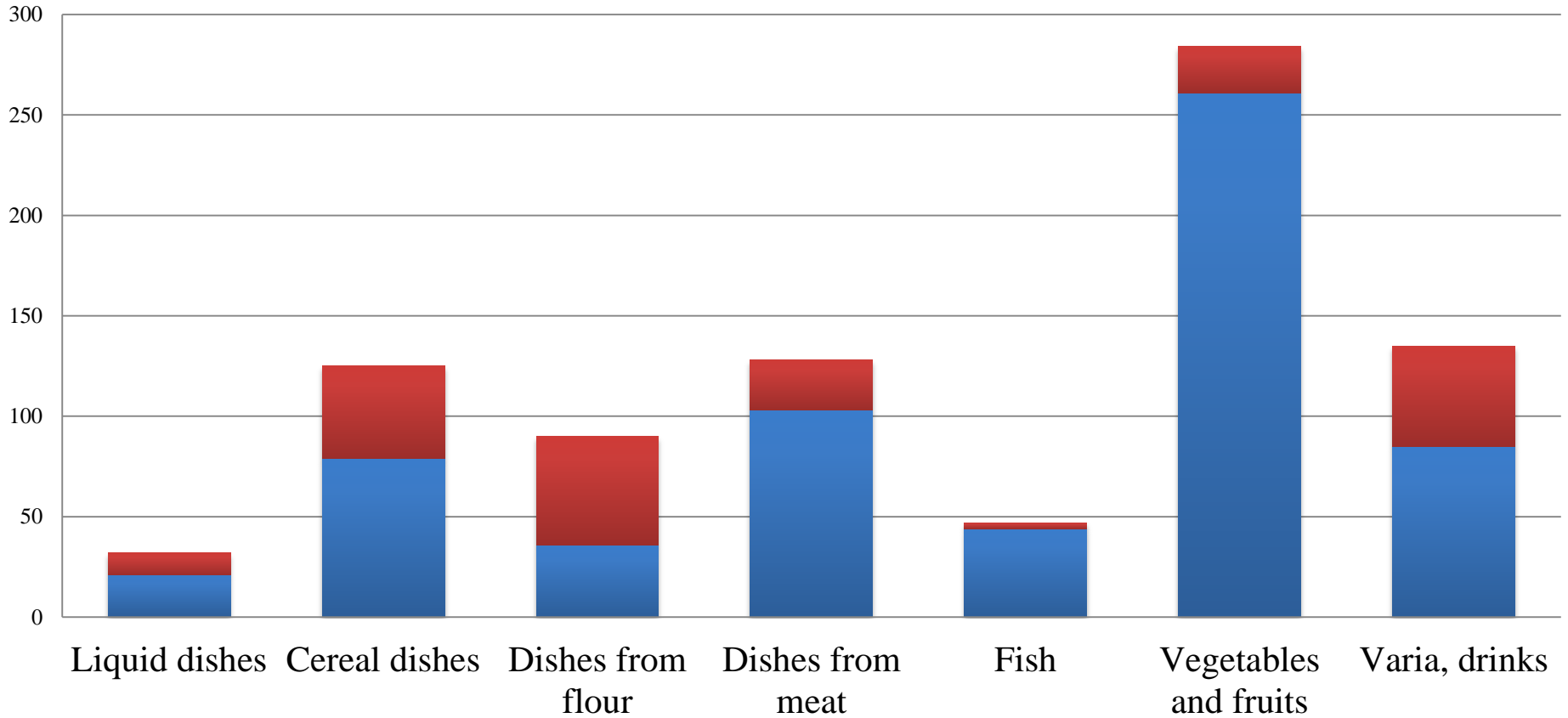
- ‘soup’ **zupa** (19) in 9 places ~ **sup** (7) in 4 places ◊  
ros./brus. суп
- ‘dumpling’ **kluski i kluseczki** (18) in 5 places ~  
**klocki i kloceczki** (30) in 13 places ◊ brus./brus.  
gw./ros. клёцки
- ‘sausage’ **kielbasa, kielbaska** (53) in 19 places ~  
**kołbasa/kałbasa, kołbaska/kałbaska** (10) in 6  
places ◊ ros. колбас|а, brus. каўбас|а
- ‘tea’ **herbata, herbatka** (30) in 15 places ~ **czaj,**  
**czajoczok** (15) in 11 places ◊ ros. чай, brus. чай i  
гарбата

## FOOD AND COOKING

- ‘fry’ *smażyć*: *posmażyć, pod smażyć, usmażyć, wysmażyć, zasmażyć* (41) ~ *podzaryć* (2) ◊ brus. см | ажыць, brus. gw. ж | арыць, ros. жарить
- ‘stew’ *dusić* (4) ~ *tuszyć i natuszyć, potuszyć, utuszyć* (14) ◊ brus. тушыць, ros. тушить
- ‘grate’ *trzeć*: *przetrzeć, zetrzeć* (12) ~ reg. wsch. *tarkować: natarkować, potarkować, utarkować* (8) ~ *darć* (2)
- ‘roll’ *wałkować: rozwałkować, powałkować* (4) ~ *kaczać, rozkaczać* (5), *cirkać* (1) ◊ brus. качаць, ros. качать, brus. цыркаць

# Proportion of native and borrowed words in synonymic pairs

■ Native words ■ Borrowings



- About 60 native and 70 borrowed nouns were analyzed, the number of which in the corpus is 76% compared to 24%.
- Among 26 noun synonymic pairs the native forms predominate over the borrowed ones (68% ~ 32%), with some exceptions.
- As for verbal pairs in the whole corpus the native forms are much more common – 87% compared to 13%.





# HOUSE AND DOMESTIC BUILDINGS



- ‘stable’ *gumno* (43), *stodoła* (36) ~ *odryna/odrynia, odryńka* (13) ◇ brus. gw. *адрына*
- ‘granary’ *spichrz/spichlerz* (5) ~ *świr, świren/świroń* (8)
- ‘shed’ *chlew* (76) in 30 places, *zasiiek* (5) in 4 places ~ *saraj, sarajczyk* (5) in 4 places
- ‘cellar’ *piwnica* (23) in 9 places ~ *sklep* (19) in 5 places ◇ SJPD daw. dziś reg. ‘loch, piwnica’, brus. gw. *склеп*
- ‘build’ *budować* (193) ~ *stroić* (153) with derivatives

# Pair of synonyms *budować* ~ *stroić* in corpus

The image shows two overlapping windows of the AntConc 3.4.4w (Windows) 2014 software. The top window displays concordance hits for the search term 'budować'. The bottom window displays concordance hits for the search term 'stroić'.

**Top Window: Concordance Hits 78**

Hit	KWIC	File
55	, ale jak poszła u n as ta, znaczy <b>budownicwa</b> tej elektrostancy, to wyrombair	KWJUF54 -3
56	en kościół ten dawny, do tamtej wojny <b>budowany</b> , bo tutaj gdzie chatka ta uo ma	KWiJI26 -71
57	tu babula jego o tuż prze tamto wojno <b>budowany</b> , to mówił szli to teraz je że te	KWiJI26 -71
58	ciężko żyć. A potym dożyli sie my, <b>budowali</b> sie, bo nie było ni budynków nic, p	KWiWM10
59	oj ja już nie pamiętam kiedy dom był <b>budowany</b> , od razu już był, tego ja nie pami	kZaFJ27, Jo:
60	ęćmi, zbiramy sie ze szkoły, tam u nich <b>budowali</b> si srebrny, czarny i taki jak srebrny	kZaFJ27, Jo:
61	, aj aj wiezie Edmuś Kukj`an doski jak <b>budował</b> mieszkanie sobie w Szarkowszczyżr	MiZK21 -31
62	ec każma moj stanisław [!], no i wot tu <b>budowla</b> ta, ot wszystko to trzeba i pole zrob	NuCS30 -7
63	sz jakich sześciemtrowa. Tak trzea było <b>budować</b> i ciongali tak j`ej. Tu chliw c	NuCS30 -7
64	yli tak. A później zbudowali dom swój, <b>budowali</b> stodole bardzo wielka u nas tutaj r	OgHS31 -7:
65	my tutaj przyjechali prosto było, sami <b>budowali</b> te mieszkania, wszystko sam pobud	PrzDrWW31
66	już już. A jak przyszli <b>sowieci normalni...</b> <b>budowali</b> sklepy. Tutaj sklepów nie dużo był	PrzDrWW31
67	siostra, i on tam w Szc	
68	am w Szczytnie swój dc	
69	oj dom <b>budował</b> , po bu	
70	w liesie mieli swój lie	

**Bottom Window: Concordance Hits 52**

Hit	KWIC	File
2	. Tak tedy kupili nowy ż domik wo, co <b>stroił</b> w siole tak, kupili ten dom on`i	ArDaWB29 -
3	j pamięci tam było, przewieźli, tutaj będo <b>stroić</b> , budować kościół ten. To ktości mów	ArDaWB29 -
4	karmi całe rodzinie, że ona tak nie umie <b>stroić</b> i tego, ali różne takie [...] ugaszcza j	ArDaWB29 -
5	ji co ?, tutaj tutaj znaczy tego, prymicja nu <b>strojim</b> , pajechali żeż do kościoła tam, nu j	ArDaWB29 -
6	czego. ja jutro nia benda w kościoła, nia <b>stroj</b> ty nic. ona mówi a czemu. Bałś	BiKG19 -L00
7	a pałoży sie, a mama piec ciepli, śniadania <b>stroji</b> , tam żywiołow trzeba papojić, pakarr	BiKG19 -L00
8	eszcze nie uczestniczyłam wtedy... Choinke <b>stroj</b> iliśmy, też tak jak umi`eliśmy, swoje za	CzWD27 -39
9	ak siak, gdzie jaka nachadziłaś rabota, i na <b>strojkach</b> przyszło sia, musi pał roku parobi	DrsTP27 -70
10	tak mama ż nie bendzie kłamać, ale czego <b>strojenka</b> choruji czego ona nie wstani. Nu	DwGrTRk30 -
11	zy zarabiali dom stawili same ten. sam tam <b>stroił</b> i sufity padbiwał, padłogi ji okny, wsz	DzJB27, DzXI
12	puszczać. Nu ja, pryd`u mamyj, mama jeść <b>stroji</b> , tam ja krowu zaprawadze, i mama ty	GawOJ36 -7
13	kwaterra, wszystko co i poszed z ojciem na <b>strojku</b> . Tylko na piontkach uczył sia, nu ja l	GaXK15, GaX
14	li i rozmajitsze, rozmajite wigilii byli. Nu co <b>strojili</b> , jak kiedy raniej, jak jeszcze mama by	GimWK31 -
15	ystwo zatrzymywał sie, jak muj tam kościół <b>stroił</b> , tam był znajomy wienkszy. A już ost	GimWK31 -
16	nie była, a tam poszła, a miała nowe <b>strojiła</b> sie, papy nie było, ojca, to stroiła sie	GimWK31 -

# CLUSTERS

***budować*** and derivatives (also ***budowa***, ***budowanie***) occur in combination with:

- ***kościół / kapliczka / parafia / szkoła / dom / mieszkanie / miasto / zamek / chlew / ściany szopa / kołowrot / droga / elektrownia (elektrostanacja)***

***stroić*** and derivatives combine with words:

- ***dom (domik) / chata / chałupy / kwatery / kościół / cerkiew / brama / łaźnia kołchoz***



While analyzing lexical material we can observe words belonging to standard language layer alongside with regional expressions (*zasiiek* i *ambar* 'storehouse', *sufit* i *stalewanie* 'ceiling', *strych* i *góra* 'attic'). These lexical units are archaic due to the disappearance of a specific type of folk culture and changes taking place in the sphere of economic activity of the countryside.

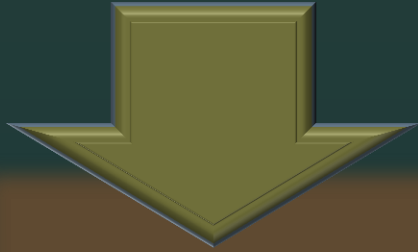


# Conclusions

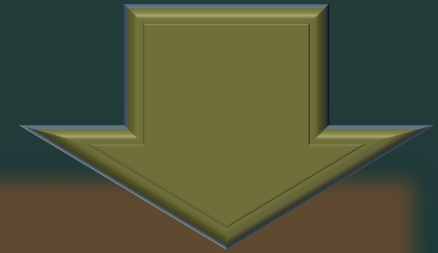
- Analyzed material clearly shows the preference of the native vocabulary in the texts, despite the cursory impression about richness of borrowings in the local Polish language.
- East Slavic words in selected pairs of synonyms often compete with Polish one, and sometimes even prevail. They affect the native system differently depending on the thematic group.



**Civilization  
transformations**



**Changing social  
and cultural conditions**



Changing names  
and their  
meanings

**VOCABULARY**

Adopting  
patterns of other  
nationalities

Continuance in  
unchanged form

**Family traditions**

