

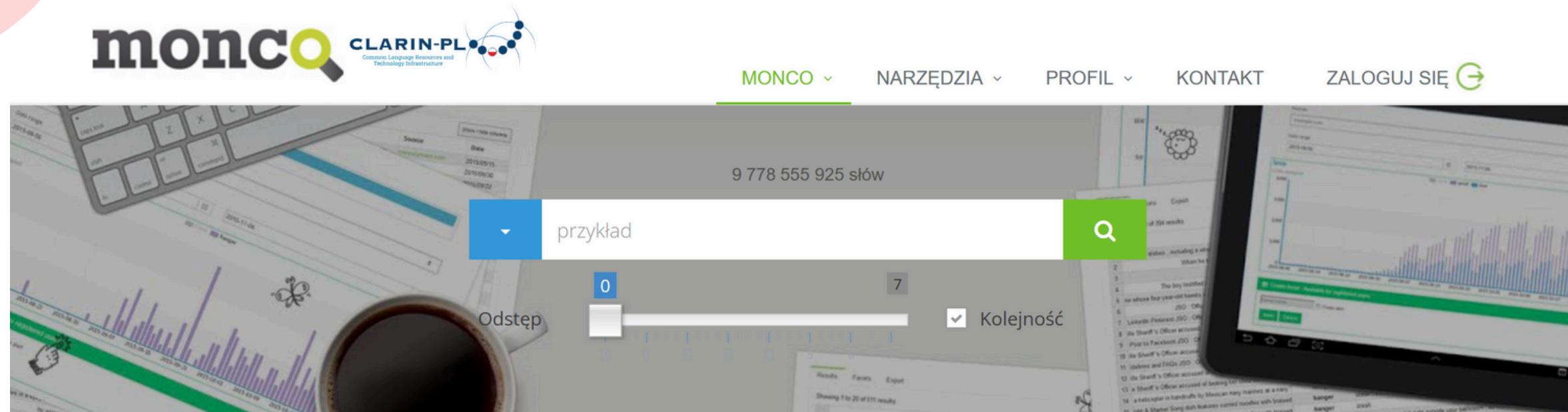
# Cyfrowe narzędzia w przekładoznawstwie

# Polskie korpusy językowe



Instrukcja dot. języka zapytań w NKJP (wyszukiwarka Poliqarp)

# Polskie korpusy językowe



O wyszukiwarce

Monco PL to wyszukiwarka korpusowa, która pomaga znaleźć przykłady użycia wyrazów, fraz oraz wzorców leksykalno-gramatycznych w autentycznych próbkach współczesnej polszczyzny.

Dane

Indeks wyszukiwarki zawiera obecnie ponad 9 miliardów segmentów słów i jest stale aktualizowany o ponad milion słów dziennie. Zaletą stałej aktualizacji jest możliwość wyszukiwania neologizmów, neosemantyzmów i zapożyczeń, które pojawiły się w polszczyźnie w ciągu kilku ostatnich lat.

# Polskie korpusy językowe

The screenshot shows the homepage of the Polish Corpus of Contemporary Polish (Korpus Współczesnego Języka Polskiego). The page has a black header bar with white text and links. On the left, there's a large red circular graphic.

**Header:**

- Korpus Współczesnego Języka Polskiego
- WYSZUKIARKA
- O KORPUSIE
- TEKSTY
- STATYSTYKI ▾
- INSTRUKCJA
- ENGLISH ▾
- LOGIN/REJESTRACJA
- Dariah.lab

**Search Form:**

Zapytanie

KONSTRUKTOR ZAPYTAŃ METADANE ▾ GRUPOWANIE WYNIKÓW ▾

Liczba wyników na stronę  
10 ▾

Wyszukaj

# Tworzenie własnych korpusów

The screenshot shows the homepage of the KORPUSOMAT website. At the top, there is a dark blue header bar with the text "KORPUSOMAT" and "NOWA WERSJA WIELOJĘZYCZNA" on the left, and navigation links "NOWY KORPUS", "MOJE KORPUSY", "DOKUMENTACJA", "KONTAKT", a language switcher (Polish flag), and "LOGIN/REJESTRACJA" on the right. Below the header is a large teal-colored main area featuring a black line-art icon of a head with a brain-like structure inside. Below the icon, the word "KORPUSOMAT" is written in large, white, sans-serif capital letters. A thin horizontal line separates this from the bottom text. At the very bottom of the teal area, the text "Utwórz własny korpus językowy" is displayed in a smaller, white, sans-serif font.

# Tworzenie własnych korpusów



AntConc

---

A freeware corpus analysis toolkit for concordancing and text analysis

[AntConc Homepage](#)

[Screenshots](#)

[Help](#)

[License](#)

[Version History](#)

# Tworzenie własnych korpusów



AntConc

---

A freeware corpus analysis toolkit for concordancing and text analysis

[AntConc Homepage](#)

[Screenshots](#)

[Help](#)

[License](#)

[Version History](#)

# Tworzenie własnych korpusów

The screenshot shows the Sketch Engine dashboard for the user 'Tyrmand'. On the left, a sidebar contains icons for different features: Word Sketch, Thesaurus, Parallel Concordance, N-grams, Trends, OneClick Dictionary, Corpus Info, Manage Corpus, Word Sketch Difference, Concordance, Wordlist, Keywords, Text type analysis, and Bilingual terms. The main area is divided into sections: 'TYRMAND' (Word Sketch, Thesaurus, Parallel Concordance, N-grams, Trends, OneClick Dictionary), 'CORPUS INFO' (Word Sketch Difference, Concordance, Wordlist, Keywords, Text type analysis, Bilingual terms), and 'RECENTLY USED CORPORA' (English Corpus for SkELL 3.10, German Corpus for SkELL 1.0, German Web 2020 (deTenTen20), Biblia). A large box on the right highlights the 'English Trends corpus' with the text: 'Explore our largest English corpus with 86+ billion words. Automatic updates: 70 million new words every week.' and a 'OPEN CORPUS' button.

DASHBOARD

Tyrmand

CORPUS INFO

MANAGE CORPUS

RECENTLY USED CORPORA

NEW CORPUS

English Corpus for SkELL 3.10

German Corpus for SkELL 1.0

German Web 2020 (deTenTen20)

Biblia

English Trends corpus

Explore our largest English corpus with 86+ billion words.

Automatic updates: 70 million new words every week.

OPEN CORPUS

Get more space +

?

!

person icon

MY SEARCH HISTORY

ANNOTATIONS

# Tworzenie własnych korpusów

The screenshot shows a search results page for 'polish' corpora. The results are listed in a table with columns for name, type, language, and size.

Name	Type	Language	Size
Polish Drama Corpus	Drama Corpus	Polish	117,230
Polish European Literary Text Collection (ELTeC)	Text Collection	Polish	8,226,827
Polish language of the 1960s	Language Corpus	Polish	546,042
Polish Parliamentary Corpus (PPC)	Parliamentary Corpus	Polish	553,858,723
Polish parliamentary debates (ParlaMint 2.1, CoNLL format)	Debate Corpus	Polish	26,882,964
Polish parliamentary debates (ParlaMint 2.1)	Debate Corpus	Polish	26,619,472
Polish Trends	Trends Corpus	Polish	1,094,089,304
Polish Web (PolishWac, Morfeusz and TaKIPi tagger)	Web Corpus	Polish	103,028,410
Polish Web 2012 (pITenTen12, RFTagger)	Web Corpus	Polish	7,715,835,214
Polish Web 2012 sample (nITenTen12)	Web Corpus	Polish	45,208,497

Below the table are three buttons: 'Show description', 'ADVANCED SEARCH', and 'CREATE CORPUS'.

# CLARIN-PL



O nas

Dla nauki

Dla biznesu

Katalog

Publikacje

Mediateka

Repozytorium

Cennik

Licencja

Aktualności

Blog

Kontakt

Polski



**CLARIN-PL**  
**Technologia języka**  
napędzająca naukę

# Inne narzędzia

The screenshot shows a web page from the University of Zurich's Cappelli Online project. At the top, there is a navigation bar with the university logo, language selection (EN), search, and login options. Below the navigation, a menu bar includes links for "Ad fontes", "Research", "Exercises", "Tutorial", and "Resources". The main content area has a breadcrumb trail ("Resources / Abbreviations /") and a title "Cappelli Online". A text block explains the nature of the crowdsourced data, mentioning errors and discrepancies. To the right, a sidebar titled "Filter Abbreviations" provides instructions on how to search for abbreviations using various fields.

All pages EN Q | ? | A-Z | Login ←

Ad fontes Research Exercises Tutorial Resources

Resources / Abbreviations /

02 / 04

Cappelli Online

The information in [Cappelli<sup>A-Z</sup>](#) online was crowdsourced and compiled with the effort of many volunteers. Though experts reviewed the work, some occasional errors remain. Spellings and thus possible discrepancies were also taken directly from the printed Cappelli into the online version. For each image of an abbreviation, you can view an image of the corresponding page in Cappelli by clicking on the page number next to the abbreviation image in the results list. Thus, the entry can be cited directly from *Ad fontes*.

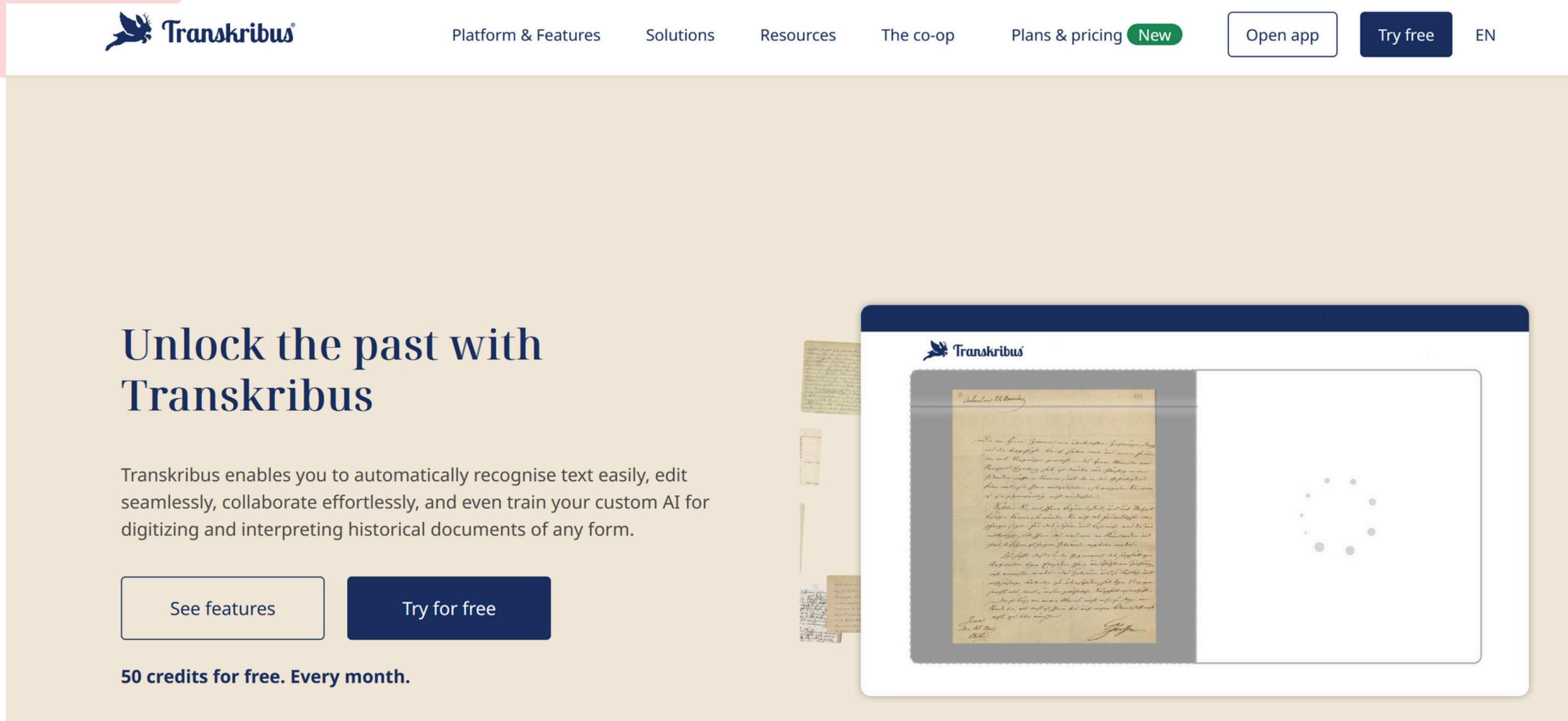
Filter Abbreviations

There are several ways to search for abbreviations. The various search fields can be used individually or in combination. Move the cursor over the question mark next to each field for more information about how to use it in your search.

# Inne narzędzia

Page ↑↓	Sign ↑↓	Transcription ↑↓	Category ↑↓	Language ↑↓	Position	
a.	1	A	Accursius, - Albericus	juristische Abkürzung	Latin	
A.	1	A	Alumen	medizinische Abkürzung	Latin	
A.	1	A	Atramentum	medizinische Abkürzung	Latin	
ꝑ.	1	A	Amen	Schrift: Westgotisch	Latin	
ꝑ	1	A	Antiphona	Schrift: Westgotisch	Latin	
ꝑ	1	A	Amalgama	medizinische Abkürzung	Latin	
ꝑ'	1	A	Antiphona		Latin	
a.	1	a	annus		Latin	

# Inne narzędzia



The screenshot shows the Transkribus website homepage. At the top, there is a navigation bar with links: Platform & Features, Solutions, Resources, The co-op, Plans & pricing (marked as New), Open app, Try free, and EN. Below the navigation bar, a large blue banner features the text "Unlock the past with Transkribus". Underneath the banner, a paragraph describes the platform's capabilities: "Transkribus enables you to automatically recognise text easily, edit seamlessly, collaborate effortlessly, and even train your custom AI for digitizing and interpreting historical documents of any form." Two buttons are present: "See features" and "Try for free". To the right of the banner, there is a preview of a historical document page with handwritten text and a transcription area. A pink circle highlights the word "Inne" in the main title "Inne narzędzia".

Transkribus

Platform & Features Solutions Resources The co-op Plans & pricing New Open app Try free EN

## Unlock the past with Transkribus

Transkribus enables you to automatically recognise text easily, edit seamlessly, collaborate effortlessly, and even train your custom AI for digitizing and interpreting historical documents of any form.

See features Try for free

50 credits for free. Every month.

Inne narzędzia

# Inne narzędzia

The screenshot shows a web browser window with the URL [www.ushuaia.pl/transliterate/](http://www.ushuaia.pl/transliterate/). The page displays the transcription of the Chinese characters "我很愛你，鬆餅" into Pinyin as "wǒ hěn ài nǐ, sōng bǐng". The interface includes a dropdown menu set to "chiński (transkrypcja pinyin)" and a "wykonaj" (execute) button. To the right, there is a sidebar with information about Hanyu Pinyin transcriptions.

rozpoznane pisma:  
chińskie

Transliterate it!

我很愛你，鬆餅

wǒ hěn ài nǐ, sōng bǐng

chiński (transkrypcja pinyin)

wykonaj

Copyright © ushuaia.pl

Transkrypcja hanyu pinyin.  
Transkrypcja standardowego języka mandaryńskiego, oficjalnie używana w Chinach, stosowana również w Polsce. Narzędzie to korzysta z bazy [unihan](#) (6.3.0).  
Więcej informacji w [wikipedia](#)  
Przykład: [天安门](#)

# Do zobaczenia w poniedziałek!

