



Narzędzia przetwarzania tekstów historycznych



Marcin Woliński, Witold Kieraś,
Dorota Komosińska, Emanuel Modrzejewski

*Zespół Inżynierii Lingwistycznej
Instytut Podstaw Informatyki
Polskiej Akademii Nauk*



✂ W Warszawie, 18 dnia Czerwca A.D. MMXVIII ✂



Przedstawione narzędzia zostały opracowane (głównie) w ramach projektu:

Model formalny diachronicznego opisu fleksji polskiej i jego komputerowa implementacja

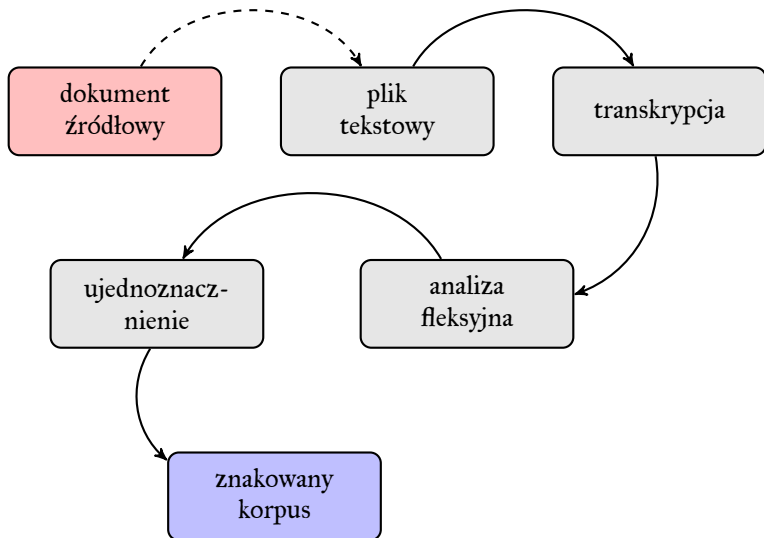
grant NCN 2014/15/B/HS2/03119

czas trwania: 11 sierpnia 2015 – 10 lutego 2019

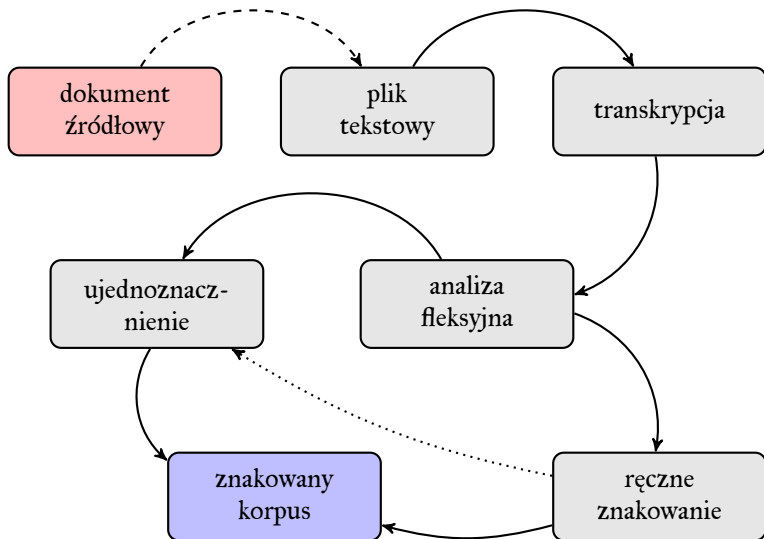
✿ Narzędzia przetwarzania tekstu barokowego ✿

- ✎ narzędzie do automatycznej transkrypcji tekstów,
- ✎ Korbeusz – automatyczny analizator fleksyjny na bazie Morfeusza 2,
- ✎ Anotatornia 2 – system do ręcznego znakowania korpusów,
- ✎ automatyczne narzędzie ujednoznaczniające (tager)

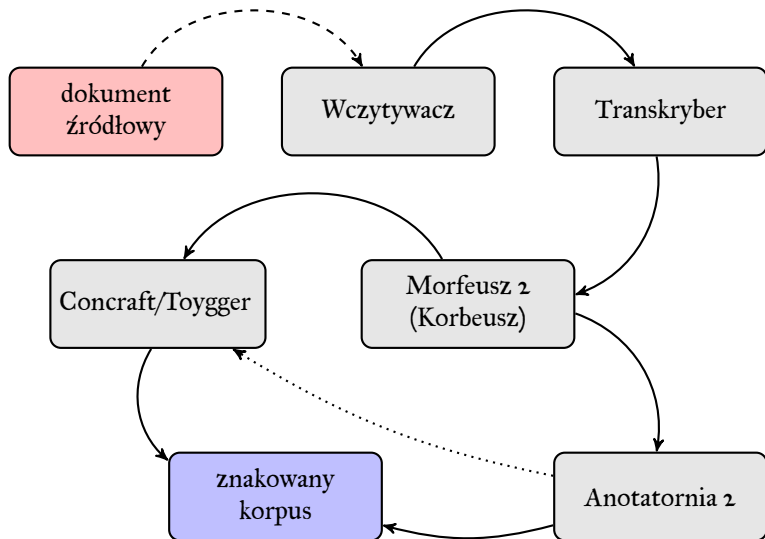
❧ Schemat działania ❧



❧ Schemat działania ❧



❧ Schemat działania ❧



Transkrypcja

-  Normalizacji podlegają wyłącznie warianty ortograficzne, nie są zaś modyfikowane dawne formy fleksyjne.
-  Nie jest to decyzja oczywista — w wielu korpusach historycznych dla innych języków normalizacji podlega ogół zjawisk językowych (czasem nawet leksyka!), by znormalizowany tekst można było przetwarzać dalej narzędziami dla współczesnego wariantu języka bez konieczności ich modyfikacji.
-  Nie zawsze niestety da się ściśle odgraniczyć zjawiska ortograficzne od fleksyjnych.

✧ Transkrypcja – przykład ✧

KOMISJA (między rokiem 1650 i 1918):

commissiey, commissyi, komisja, komisją, komisję, komisji, komissja, komissya, komissyę, komissyi, komisya, komisją, komisye, komisję, komisyi, komisyj, komisją, komisję, kommissij, kommissja, kommissjów, kommissii, kommissij, kommissją, kommissje, kommissjów, kommissya, kommissyi, kommissyiey, kommissyi.

W procesie transkrypcji:

-  zmieniono inicjalne *c* na *k*,
-  usunięto podwojenia liter spółgłoskowych *s* i *m*,
-  wprowadzono *j* zamiast *y* lub *i*.
-  pozostawiono dawne formy fleksyjne: *komisjej*, *komisjów* i *komisją* (w B. lp.).

Transkrypcja – przykłady reguł

	lewy kontekst	znajdź	prawy kontekst	zmień na	wyjątki
1.	.*	$\hat{o}\hat{o}$	.*	go	–
2.	$\wedge$	naiw	.*	najw	naiwn.*
3.	T	j	T	y	mjr

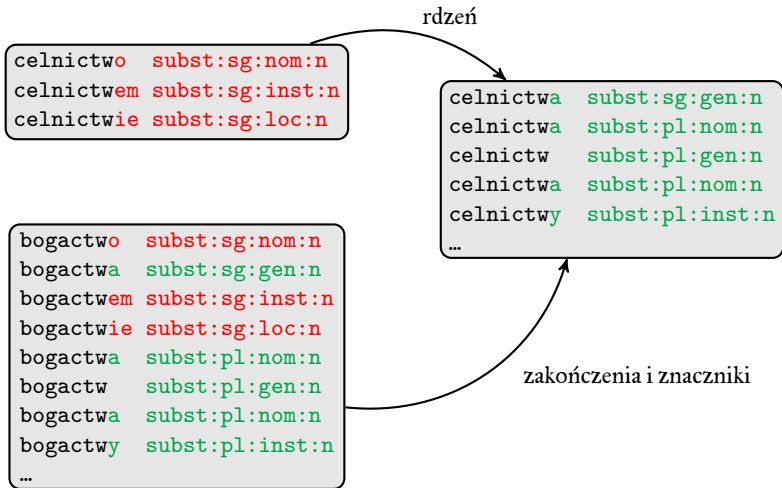
Słownik:

- ✚ „postarzone” dane SGJP — konwersja tagsetu i dotworzenie niektórych regularnych dawnych form (*walidaczej, ładną, ładnemi, zmienim, klęczta, klęczwa* itp.),
- ✚ dane fleksyjne SXVII,
- ✚ automatycznie zrekonstruowane częściowe paradygmaty SXVII.

Reguły segmentacyjne:

- ✚ pisownia łączna z *nie*,
- ✚ pisownia łączna licznych partykuł (*ć, li, że, ż*),
- ✚ znacznie bardziej ruchome aglutynanty.

Automatyczna rekonstrukcja



Tagery



Concraft:

- * conditional random fields,
- * ujednoznacznianie niejednoznacznej segmentacji (94%),
- * słabsze wyniki dezambiguacji (88%),
- * nieco lepsze zgadywanie.



Toygger:

- * sieci neuronowe (bi-LSTM),
- * wyłącznie ujednoznacznianie znaczników na zadanej segmentacji,
- * lepsze wyniki dezambiguacji (91%),
- * nieco słabsze zgadywanie.

✧ Anotatoria 2 ✧



System Anotatoria 2 służy do ręcznego ujednolicenia i weryfikacji znakowania fleksyjnego tekstu historycznego.



Uwzględnia specyficzne potrzeby korpusów historycznych:

- ✧ tekst istnieje w dwóch równoległych postaciach: transliterowanej i transkrybowanej,
- ✧ musi być zachowywana informacja o numerach stron oryginałów, z których pochodzą poszczególne wyrazy.

<http://zil.ipipan.waw.pl/Anotatoria2>

✧ Anotatoria 2 ✧

Praca w Anotatorni obejmuje:

-  korektę segmentacji – wyróżnienia wykładników tekstowych form fleksyjnych,
-  korektę transkrypcji,
-  korektę podziału na zdania,
-  ujednolicenie i uzupełnienie opisu fleksyjnego dostarczonego przez Korbeusza.

⌘ Anotatornia 2: Tryb znakowania AA+A ⌘

Powszechnie przyjęty tryb znakowania korpusów:

- ☞ każda próbka korpusowa jest znakowana niezależnie przez dwóch annotatorów,
- ☞ lista konfliktów jest przedstawiana annotatorom do ponownego sprawdzenia,
- ☞ pozostałe konflikty rozstrzyga arbiter („superannotator”).

Pomysł usprawnienia: zastąpić jednego z annotatorów programem automatycznym (tagerem):

- ☞ każda próbka korpusowa jest znakowana tylko przez jednego annotatora,
- ☞ wynik jego pracy jest porównywany z wynikami tagera,
- ☞ lista konfliktów jest przedstawiana annotatorowi do ponownego sprawdzenia,
- ☞ pozostałe konflikty rozstrzyga arbiter.

Segmenty



soluriui| vnc|.↵

a conj



nie qub

- ustawać praet pl:manim2:imperf



- ustawać praet pl:mnanim:imperf



SykstCiepl, próbka 1

Uwagi

Gotowe

Tekst transliterowany

A| gdzie|by| wymioty| nie| **vstawály** ;| daley| iuż| nie|
postępuy| aby|ś| miał| roskázować| tę| wodę| pić|:| ále|
myśl| o| **czy|m|** **czym**| innym| iákoby|ś| zdrowie| onego|
pacyentá| rátował|.↵

Drugi| przypadek| iest|/| kiedy| wodá| zostáie| się| w|
ćiele|/| y| nie| wychodzi|:| á| zostáie| ábo| w| ielitách|/| abo|
też| we| wszystkim| ćiele| y| w| **żyłá|ch|** **żyłách**|.↵

Ná| to| powiedział|em| ia| iuż| naukę| iednę|:| iż| kiedy|
chcemy| áby| przez| stolce| tá| wodá| wychodziłá|/| day|
mu| w| przod| ná| każdy| dzień| pierwey| **niże|li|** **niżeli**|
pocźnie| pić| wodę| ábo| wánny| vnc|.↵

ii|.↵

ábo| też| syr|.↵

rosat|.↵

soluriui| vnc|.↵

Segmenty

A

a conj

gdzie

gdzie adv

by

by qub

wymioty

wymioty subst pl:nom:p2

nie

nie qub

ustawály

• ustawać praet pl:manim2:imperf

• ustawać praet pl:mnanim:imperf

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe ✓

Tekst transliterowany

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człeczka| tu| wykroczy|,| Na| świat|,
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵Jeśli| jakie| przewiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵Nie| wzdryga| się|,| imiona| przewiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzi|.↵Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| tem|,
Niepotrzebnym| się| **cale** zarażać| kłopotem|.↵Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| Kiedy|ś|
Kiedyś| szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyńfy|.↵

Niepotrzebnym

niepotrzebny
adj sg:inst:m:pos

się

się qub

cale

- cal subst pl:nom:manim2 ✎
- cal subst pl:nom:m ✎
- cal subst pl:acc:manim2 ✎
- cal subst pl:acc:m ✎
- cal subst pl:voc:manim2 ✎
- cal subst pl:voc:m ✎
- cale adv ✎

zarażać

- zarażać inf imper ✎

kłopotem

- kłopot subst sg:inst:m ✎

?

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe

Tekst transliterowany

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,|
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵

Jeśli| jakie| przewiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵

Nie| wzdryga| się|,| imiona| przewiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵

Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| tem|,|
Niepotrzebnym| się| **cale** zarażać| kłopotem|.↵

Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| Kiedy|ś|
Kiedyś| szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵

Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵

Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyńfy|.↵

Niepotrzebnym

niepotrzebny
adj sg:inst:m:pos

się

się qub

cale

cale adv

zarażać

• zarażać inf imperf

kłopotem

• kłopot subst sg:inst:m

?

? interp

Albo

• albo qub

ś

PotFrasz2Kuk II, próbka 1

Uwagi

Gotowe

Tekst transliterowany

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,|
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵

Jeśli| jakie| przezwiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵

Nie| wzdryga| się|,| imiona| przezwiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵

Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| **tem**|,|
Niepotrzebnym| się| całe| zarażać| kłopotem|.↵

Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| Kiedy|ś|
Kiedyś| szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵

Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵

Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tynty|.↵

tez qub

było

być praet sg:n:imperf

po

po prep loc

tem

- tema subst pl:gen:f
- ten adj sg:inst:m:pos
- ten adj sg:inst:n:pos
- ten adj sg:loc:m:pos
- ten adj sg:loc:n:pos

,

, interp

Niepotrzebnym

- niepotrzebny adj sg:inst:m:pos
- niepotrzebny adj sg:inst:n:pos
- niepotrzebny adj sg:loc:m:pos

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe ✓

↔️ Tekst transliterowany

⏮️ ⏭️

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵Jeśli| jakie| przewiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵Nie| wzdryga| się|,| imiona| przewiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| **tem**|,
Niepotrzebnym| się| całe| zaraża| kłopotem|.↵Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| Kiedy|ś|
Kiedyś| szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyńfy|.↵

tez qub

było

A 🗑️ 🔗 ↵

być praet sg:n:imperf

po

A 🗑️ 🔗 ↵

po prep loc

tem

A 🗑️ 🔗 ↵

Lemat:

to

Cz.m.:

subst

Tag:

sg:loc:n

OK ✓

Anuluj ✖️

• tema subst pl:gen:f

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe ✓

↔️ Tekst transliterowany

⏮️ ⏭️

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,|
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵Jeśli| jakie| przewiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵Nie| wzdryga| się|,| imiona| przewiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| **tem**|,|
Niepotrzebnym| się| całe| zarażać| kłopotem|.↵Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| Kiedy|ś|
Kiedyś| szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyńfy|.↵

tez qub

było

być praet sg:n:imperf

po

po prep loc

tem

✎ to subst sg:loc:n

,

, interp

Niepotrzebnym• niepotrzebny
adj sg:inst:m:pos ✎

• niepotrzebny adj sg:inst:n:pos

✎

• niepotrzebny adj sg:loc:m:pos

✎

• niepotrzebny adj sg:loc:n:pos

✎

• niepotrzebny adj pl:dat:m:pos

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe

Tekst transliterowany

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵

Jeśli| jakie| przezwiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵

Nie| wzdryga| się|,| imiona| przezwiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵

Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| tem|,
Niepotrzebnym| się| cale| zarażać| kłopotem|.↵

Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| **Kiedy** |ś|
Kiedys |szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵

Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵

Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyńfy|.↵

upić praet sg:m:perf

,

, interp

Kiedy

- kiedy adv
- kiedy comp
- kiedy qub

ś

- być aglt sg:sec:imperf:nwok

Kiedys

- kiedys adv

szyderstwo

- szyderstwo subst sg:nom:n
- szyderstwo subst sg:acc:n
- szyderstwo subst sg:voc:n

z

- z prep gen:nwok

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe ✓

Tekst transliterowany

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,|
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵Jeśli| jakie| przewiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵Nie| wzdryga| się|,| imiona| przewiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| tem|,|
Niepotrzebnym| się| cale| zarażać| kłopotem|.↵Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| **Kiedys**
szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|.↵Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyńfy|.↵

upić praet sg:m:perf

,

, interp

Kiedys

kiedys adv

szyderstwo

- szyderstwo subst sg:nom:n
- szyderstwo subst sg:acc:n
- szyderstwo subst sg:voc:n

z

- z prep gen:nwok
- z prep acc:nwok
- z prep inst:nwok
- zeszyt brev pun

siebie

- siebie siebie gen
- siebie siebie acc

PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe

Tekst transliterowany

jeden| grzech|,| lecz| każdy| inszy| we| mnie| czyta|.↵

Nie| człek|,| ale| natura| człecz| tu| wykroczy|,| Na| świat|,
nie| w| się| mu| w| głowie| obróciwszy| oczy|.↵

Jeśli| jakie| przewiska|,| takie| i| humory|,| Morstyn| sobie|
Oborska| przybiera| do| sfory|.↵

Nie| wzdryga| się|,| imiona| przewiska| łagodzą|:
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|.↵

Pytam|,| mój| miły| bracie|,| co|ć| też| było| po| tem|,
Niepotrzebnym| się| cale| zarażać| kłopotem|?↵

Albo|ś| szalał|,| albo|ś| się| natenczas| był| upił|,| Kiedyś|
szyderstwo| z| siebie| w| tym| urzędzie| kupił|.↵

Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić|,| Gdy|
trzeba|,| choć| kto| błaznem| chce| zostać|,| zapłacić|?↵

Leśne| dryjady| i| wy|,| morskie| wiedzcie| nimfy|,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyfny|.↵

tej

A ↵ ↶ ↷

ten adj sg:gen:f:pos

ligi

A ↵ ↶ ↷

liga subst sg:gen:f

wchodzą

A ↵ ↶ ↷

wchodzić fin pl:ter:imperf

.

A ↵ ↶ ↷ ↵

. interp

Pytam

A ↵ ↶ ↷

pytać fin sg:pri:imperf

,

A ↵ ↶ ↷

, interp

mój

A ↵ ↶ ↷

mój subst sg:voc:m

miły

A ↵ ↶ ↷

← PotFrasz2Kuk_II, próbka 1

Uwagi

Gotowe ✓

Tekst transliterowany

Podobno,| jeśli o| mnie| kto| wszystkich| tych| spyta,| Nie|
jeden| grzech,| lecz| każdy| inszy| we| mnie| czyta.↵

Nie| człek,| ale| natura| człeczka| tu| wykroczy,| Na| świat,|
nie| w| się| mu| w| głowie| obróciwszy| czyi.↵

Jeśli| jakie| przezwiska,| takie| i| humory,| Morstyn| sobie|
Oborska| przybiera| do| sfory.↵

Nie| wzdygaj| się,| imiona| przezwiska| łagodzą:|
Stanisław| z| Konstancją| do| tej| ligi| wchodzą|
Pytam,| mój| miły| bracie,| co| też| było| po| tem,|
Niepotrzebnym| się| cale| zarażać| kłopotem?↵

Albo|ś| szalał,| albo|ś| się| natenczas| był| upił,| Kiedyś|
szyderstwo| z| siebie| w| tym| urzędzie| kupił.↵

Na| kiego|ż| kata| w| Polszcze| pieniędzy| nie| tracić,| Gdy|
trzeba,| choć| kto| błaznem| chce| zostać,| zapłacić?↵

Leśne| dryjady| i| wy,| morskie| wiedzcie| nimfy,| Że| ten|
pan| dzisiaj| został| kpem| za| swoje| tyfny.↵

tej

A

↵

↵

↵

ten adj sg:gen:f:pos

ligi

A

↵

↵

↵

liga subst sg:gen:f

połącz segment z następnym

wchodzą

A

↵

↵

↵

wchodzić fin pl:ter:imperf

.

A

↵

↵

↵

. interp

Pytam

A

↵

↵

↵

pytać fin sg:pri:imperf

,

A

↵

↵

↵

, interp

mój

A

↵

↵

↵

mój subst sg:voc:m

miły

A

↵

↵

↵

RelWjazdKról, próbka 1

Uwagi

Gotowe

Tekst transliterowany

Piechoty| Krolewskie| á| samym| wieczorem|,| były|
 Fáierwetki| przed| Ratuszem|,| dość| misterne|,| bo| ná|

Powietrzu| wydawały|
 gorzały|,| to| iest|,| viv|
 wypadáły|.↵

Muzyki| wszelákie| gr|
 Laudamus| śpiewano|

W| Rynku| cztery| wo|
 kuf| wielkich| winá|,

wystáwiono|,| ktore| miáto| z| fontány| wypadáć|,| dla|
 krotkości| czasu| nie| wygotowáney|,| etc|.| etc|.↵

Owo| zgoła|,| krotkiami| słowy| przyidzie| záwrzeć| tę|
 Relátya|,| że| áni| Wiázdu| ták| Solennego|,| y| wiekszą|
 Pompá| odpráwionego| saecula| nie| miály|,| choćy| też| y|
 trzech| Monárchow| insimul| Wiázdy| swoje| czyniły|,|
 wszystko| spokojnie| z| łáski| Bożey| odpráwiło| się|,| Ad|
 Potentissimi| Regis| Poloniarum| Honorem|.↵

WYKRYTO KOLIZJE

i

A

i conj

kilka

A

kilka num pl:acc:f

A

f

A

acc:f:pos

A

wino subst sg:gen:n

A

, interp

ex <foreign>

A

Mandato <foreign>

A

RelWjazdKról, próbka 1 ➡

Uwagi 📎

Gotowe ✓

Tekst transliterowany



...
 Piechoty| Krolewskie| á| samym| wieczorem|,| byli|
 Fáierwetki| przed| Ratuszem|,| dość| misterne|,| bo| ná|
 Powietrzu| wydawały| igrzyská| rozmaíte|,| y| pewne| litery|
 gorzały|,| to| iest| vivat| AVGVSTVS| etc|,| y| ráce| z| nich|
 wypadáły|.

Muzyki| wszelákie| grály| ná| Ratuszu|,| Te| Deum|
 Laudamus| śpiewano|,| etc|.

W| Rynku| cztery| woły| nádziejáne| pieczono|,| y| kilká|
 kuf| wielkich| winá|,| ex| Mandato| Regio| pro| Po|puło|
 wystáwiono|,| ktore| miáło| z| fontány| wypadáć|,| dla|
 krotkości| času| nie| wygotowáney|,| etc|,| etc|.

Owo| zgoła|,| krotkiami| słowy| przyidzie| záwrzeć| tę|
 Relátya|,| że| áni| Wiázdu| ták| Solennego|,| y| wiekszą|
 Pompą| odpráwionego| saecula| nie| miały|,| choćy| też| y|
 trzech| Monárchow| insimul| Wiázdy| swoje| czyniły|,|
 wszystko| spokojnie| z| łáski| Bożey| odpráwiło| się|,| Ad|
 Potentissimi| Regis| Poloniarum| Honorem|.

i



i conj

kilka



kilka num pl:acc:f

kuf



kufa subst pl:gen:f

wielkich



wielki adj pl:acc:f:pos

wina



wino subst sg:gen:n

,



, interp

ex <foreign>



Mandato <foreign>



SulListyKoniec, próbka 1

Gotowe ✓

Tekst transliterowany

niedawno| do| nieboszczyka| Dilawer-basze| Dilawer-basze| Dilawer-basze| pisany|,| z| powinności| moiey| pilno| sie| stara|,| o| to|,| żeb| by|,| by|,| by|,| Baszowie|

ze|
prz|
z|
Jg|
zd|
gd|
Po|

przed| tymi| tymi| tymi| dwiema| namiestnikami| tego|,| a| zostawował| tu| Husejn-basze| Husejn-basze| Husejn-basze| i| Giurdzi-Machmet - basze| Giurdzi-Machmet - basze| Giurdzi-Machmet| - basze|,| albo| tesz| za| nim| do| Damaszku| iechał|,| sam| P|,| Bóg| wyuzdanemu| animuszowi| na| dishonor| dishonor| dishonor| J| J| J|,| Kr| Kr|,| Mści| Mści| Mści| y| winę| Rzpłtey| przybrał| munsztuk| z| łaski| y| opatrności| swoiey|,| y| mniej| sam| nie|lada| nie|lada| nie|lada| jako| ucieszył|,| podawszy|

sprawować praet sg:m:imperf

przed

przed prep inst:nwok

tymi

ten adj pl:inst:m:pos

INTERFEJS SUPERANNOTATORA

- ten adj pl:inst:n:pos
- ten adj pl:inst:p1:pos
- ten adj pl:inst:p2:pos

dwiema

dwa num du:inst:m

namiestnikami

namiestnik subst pl:inst:m

SulListyKoniec, próbka 1

Gotowe ✓

Tekst transliterowany

niedawno| do| nieboszczyka| Dilawer-basze| Dilawer-basze| Dilawer-basze| pisany|,| z| powinności| moiey| pilno|m| się| starał| o| to|,| żeby| byli| byli| byli| Baszowie| zeszłego| zeszłego| zeszłego| Cesarza| zatrzymać| do| przyjazdu| J| J| J|. Mści| Mści| Mści| Pana| Posła| mogli|,| ale| z| upornym| trudno| było| co| począć|,| który| w| sprawie| Jg^o^A| Jg^o^A| Jg^o^A| Kr| Kr| Kr|. Mości| lekkomyślnie| nad| zdanie| rady| swey| taką| był| ordynatę| uczynił|,| że| gdyby| był| pod| niebytność| iego| J| J| J|. M| M| M|. P|. Poseł| przyjechał|,| chciał| żeby| poselstwo| sprawował| przed| tymi| tymi| tymi| dwiema| namiestnikami| iego|,| a| zostawował| tu| Husejn-basze| Husejn-basze| Husejn-basze| i| Giurdzi-Machmet| -| basze| Giurdzi-Machmet-basze| Giurdzi-Machmet| -| basze|,| albo| tesz| za| nim| do| Damaszku| iechał|,| sam| P|. Bóg| wyuzdanemu| animuszowi| na| dishonor| dishonor| dishonor| J| J| J|. Kr| Kr| Kr|. Mści| Mści| Mści| y| winę| Rzpłtey| przybrał| munsztuk| z| łaski| y| opatrności| swoiey|,| y| mniej| sam| nie|lada| nie|lada| nie|lada| jako| ucieszył|,| podawszy|

sprawować praet sg:m:imperf

przed

przed prep inst:nwok

tymi

ten adj pl:inst:m:pos

tymi

ten adj du:inst:m:pos

tymi

- ten adj pl:inst:m:pos
- ten adj pl:inst:f:pos
- ten adj pl:inst:n:pos
- ten adj pl:inst:p1:pos
- ten adj pl:inst:p2:pos

dwiema

dwa num du:inst:m

namiestnikami

namiestnik subst pl:inst:m

SulListyKoniec, próbka 1

Gotowe ✓

Tekst transliterowany

zdanie| rady| swey| taką| był| ordynację| uczynił|,| że|
 gdyby| był| pod| niebytność| iego| J J J J. | M M M M. | P. |
 Poseł| przyjechał|,| chciał| żeby| poselstwo| sprawował|
 przed| tymi| tymi| tymi| dwiema| namiestnikami| iego|,| a|
 zostawował| tu| Husejn-baszę| Husejn-baszę| Husejn-
 baszę| i| Giurdzi-Machmet| -| baszę| Giurdzi-Machmet -
 baszę| Giurdzi-Machmet| -| baszę|,| albo| tesz| za| nim| do|
 Damaszku| iechał|;| sam| P. | Bóg| wyuzdanemu|
 animuszowi| na| dishonor| dishonor| dishonor| J J J J. | Kr
 Kr Kr. | Mści| Mści| Mści| y| winę| Rzpltey| przybrał|
 munsztuk| z| łaski| y| opatrności| swoiey|,| y| mnie| sam|
 nie|lada| nielada| nielada| jako| ucieszył|,| podawszy|
 nachyloną| do| confusiey| sprawę| pokoju| Rzpltey| w| ręce|
 terazniejszemu| Cesarzowi|,| Panu| cnot| z| przyrodzenia|
 przyrodzenia| przyrodzenia| wielkich|,| który| usiadłszy| na|
 stolicy|,| z| tym| się| naprzód| ozwał|,| że| ze| wszystkimi|
 Pany| Chrześcianańskimi| sąmsiady| swymi|,| którzy|by|
 tego| wdzięczni| bydź| chcieli| y| osobiwie| z| J J J J. | K K
 K. | Mcią| Mcią| Mcią| chce| y| życzy| sobie| pokoju|.↵

mnie

ja ppron12 sg:acc:m:pri:akc

sam

sam adj sg:nom:m:pos

nie

nie qub

lada

lada qub

nielada

nielada adv

nielada

nielada subst sg:nom:f

jako

jako adv

ucieszył

ucieszyć praet sg:m:perf

✧ Ręcznie oznakowana część korpusu ✧



W ramach projektu KorBa oznakowano w Anotatorni 2 fragment korpusu barokowego:

- * w trybie AA+A zostało oznakowane 365 514 segmentów,
- * ponadto 146 240 segmentów zostało oznakowanych tylko przez jednego annotatora.
- * Łączna długość korpusu znakowanego ręcznie: 511 754 segmenty.

Podsumowanie



Narzędzia opracowane w ramach projektu Chronofleks (a więc przede wszystkim odpowiednie wersje Morfeusza 2 i Anotatoria 2) zostały użyte do znakowania korpusów z trzech okresów:

- * barokowego,
- * XIX wieku,
- * współczesnego.



Pozwala to wyrazić przekonanie, że narzędzia te można zastosować również do tekstów z innych epok.

Do czego serdecznie zachęcamy.