



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

Zasób leksykalny polszczyzny II poł. XIX wieku a możliwość automatycznej analizy morfologicznej tekstów z tego okresu

DEC-2012/07/B/HS2/00570

Magdalena Derwojedowa Witold Kieraś
Danuta Skowrońska Robert Wołosz

IJP UW, Uniwersytet w Pécsu



Synchronia i diachronia, Toruń, 5–6 września 2013, Toruń

Porządek wystąpienia



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- 1 Projekt
- 2 Narzędzia do przetwarzania tekstów języka polskiego
- 3 Automatyczna analiza morfologiczna
- 4 *Słownik języka polskiego* pod red. W. Doroszewskiego
- 5 Materiał
- 6 Testy
- 7 Podsumowanie: synchronia i diachronia

Projekt



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

Projekt *Automatyczna analiza fleksyjna tekstów polskich z lat 1830–1918 z uwzględnieniem zmian w odmianie i pisowni 2013–2016, DEC-2012/07/B/HS2/00570.*

Cele:

- koncepcja leksykograficzna opisu zmian fleksyjnych i pisownianych w (elektronicznym) słowniku gramatycznym
- niewielki, zrównoważony i oznakowany (1 mln segmentów) korpusu tekstów z lat 1830–1918
- wzorce paradygmatyczne nieuwzględnione w SGJP, żywe w badanym okresie
- systematyka kategorii fleksyjnych i ich wartości
- schemat opisu zjawisk ewolucyjnych w słowniku elektronicznym
- poszerzenie słownika analizatora morfologicznego

Narzędzia do przetwarzania tekstów języka polskiego



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- synchronia: zasoby dla współczesnej polszczyzny
- diachronia: komputerowe słowniki i programy do pracy z tekstami dawnymi

Synchronia: zasoby



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- korpusy (IPI PAN, PELCRA, NKJP,
<http://http://korpus.pl/>, <http://pelcra.pl/>,
<http://www.nkjp.pl/>)
- zasoby semantyczno-leksykalne (polski wordnet, *Słowosieć*,
<http://plwordnet.pwr.wroc.pl/wordnet/>)
- bank drzew (*Składnica*,
<http://zil.ipipan.waw.pl/Sk%C5%82adnica/>)
- słowniki walencyjne (*Walenty*,
<http://clip.ipipan.waw.pl/Walenty>)

Synchronia: narzędzia



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- analizatory morfologiczne (Morfeusz, Polimorf, PolEng, PoMor, AMOR)
- analizatory składniowe (Świga,
<http://nlp.ipipan.waw.pl/~wolinski/swigra/>,
analizator T. Obrębskiego, <http://www.staff.amu.edu.pl/~obrebski/files/to-dr/to-dr.pdf>)
- zestawy narzędzi do pracy z tekstem (Poliqarp, PSI toolkit, narzędzia grupy G4.19,
<http://poliqarp.sourceforge.net/>,
<http://psi-toolkit.wmi.amu.edu.pl/>,
<http://nlp.pwr.wroc.pl/pl/narzedzia-i-zasoby/narzedzia-przetwarzania-morfosyntaktycznego>)

Analizatory morfologiczne



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- Morfeusz, <http://sgjp.pl/morfeusz/>
- Polimorf, <http://zil.ipipan.waw.pl/PoliMorf>
- Słownik fleksyjny języka polskiego na CD (<http://www.klk.uj.edu.pl/sfjp>)
- system POLENG (<http://poleng.pl/>)
- analizator SAM
- analizator PoMor
- analizator AMOR

Diachronia: zasoby i programy



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- *Korpus tekstów staropolskich* (<http://www.ijp-pan.krakow.pl/publikacje-elektroniczne/korpus-tekstow-staropolskich>)
- *Elektroniczny korpus tekstów polskich z XVII i XVIII w. (do roku 1772)* (grant 11H 12 0124 81 realizowany w Instytucie Języka Polskiego PAN, kierownik dr hab., prof. SWPS Włodzimierz Gruszczyński)
- *Słownik polszczyzny XVI w.* (<http://www.spxvi.edu.pl/>)
- *Słownik polszczyzny XVII i 1 poł. XVIII wieku* (http://xvii-wiek.ijp-pan.krakow.pl/pan_klient/)
- *Evangelie*

Automatyczna analiza morfologiczna



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

**Automatyczna analiza
morfologiczna**

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

Analiza morfologiczna

operacja przyporządkowująca każdemu słowu wejściowemu pewien opis (np. poprawność ortograficzną, postać hasłową, opis fleksyjny, słowotwórczy itd.).

Analiza tekstów dawniejszych



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie: synchronia i diachronia

- Analizator o rozległej podstawie leksykalnej może poprawnie analizować znaczący odsetek jednostek dawniejszych.

Analiza tekstów dawniejszych



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie: synchronia i diachronia

- Analizator o rozległej podstawie leksykalnej może poprawnie analizować znaczący odsetek jednostek dawniejszych.
- Tworzenie nowego analizatora jest niecelowe.

Analiza tekstów dawniejszych



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie: synchronia i diachronia

- Analizator o rozległej podstawie leksykalnej może poprawnie analizować znaczący odsetek jednostek dawniejszych.
- Tworzenie nowego analizatora jest niecelowe.
- Modyfikacja słownika istniejącego bogatego analizatora.

Słownik języka polskiego

pod red. W. Doroszewskiego



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

liczba haseł

ok. 125 000

liczba haseł oznaczonych jako *dawne*

10 597

liczba haseł oznaczonych jako *przestarzałe*

4 606

15 203

Słownik języka polskiego pod red. W. Doroszewskiego



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

liczba haseł ok. 125 000

liczba haseł oznaczonych jako *dawne* 10 597

liczba haseł oznaczonych jako *przestarzałe* 4606

15 203

ok.11,5% ogólnej liczby haseł

Hasła *dawne* i *przestarzałe* w SJPDor



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

biegus *m IV, Ms.* ~ *usie 1. zool.* «Tringa can-
ctus, ptak z rodziny bekasów, żyje w Europie,
Azji i Ameryce na obszarach północnych»:
Grały bekasy, chruściele, łyski krzykliwe i bie-
gusy błotne. *KONOPN. Now. 11, 272.*

2. daw. «stworzenie, które biega, zwłaszcza
koń»: Twardowski dosiadł biegusa, próbuje
podskoków, zwrotów, stępa, galopuje, kłusa.
MICK. Ball. 50. Droga zakręcała się w prawo
pod górę, woźnica zwolnił cugła biegusom.
SIERP. S. Pow. 11, 21. // L

Hasła *dawne* i *przestarzałe* w SJPDor



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

biegus *m IV, Ms.* ~ *usie 1. zool.* «Tringa can-
ctus, ptak z rodziny bekasów, żyje w Europie,
Azji i Ameryce na obszarach północnych»:
Grały bekasy, chruściele, łyski krzykliwe i bie-
gusy błotne. *KONOPN. Now. 11, 272.*

2. daw. «stworzenie, które biega, zwłaszcza
koń»: Twardowski dosiadł biegusa, próbuje
podskoków, zwrotów, stępa, galopuje, kłusa.
MICK. Ball. 50. Droga zakręcała się w prawo
pod górę, woźnica zwolnił cugła biegusom.
SIERP. S. Pow. 11, 21. // L

kraszuarka *ż III, Im D.* ~ *rek przestarz.*
«spluwaczka»: W drewnianej kraszuarce wię-
cej było płwocin niż piasku. *ZER. Syzyf. 93.*
// SWil
(*fr. crachoir*)

Materiał



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- (1) W o g ó l e żaden na świecie język nie jest mi drogim i nie ma w moich oczach żadnych praw.
- (2) D r o g i e m mi jest prawo człowieka pozostawania przy własnym języku, wybierania go sobie, prawo n i e p o d l e g a n i a wywłaszczaniu ze wszechstronnej używalności swego języka, prawo dowolnego grupowania się ludzi między innymi także na podstawie językowej.
- (3) Krótka h i s t o r y a idei języka międzynarodowego i jej wcieleń.
- (4) Do usunięcia wielojęzyczności dążono r ó ż n e m i s p o s o b y, drogą krzywd i gwałtów.
- (5) módz, biedz
- (6) kategorii, biblioteczka

Testy



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- *Lalka* B. Prusa, 260 000 segmentów
- *Pan Tadeusz* A. Mickiewicza
- wybrane pisma J. N. Baudouin de Courtenay
- analizatory PoMor i Morfeusz-SGJP



Nierozpoznane segmenty i ich charakterystyka

Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

Morfeusz SGJP, PoMor

nierozpoznane formy
4000

Nierozpoznane, charakterystyka



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- elementy obce (liczby itp.)
- nazwy własne:
 - (7) Założyłbym się, że **K l e j n** będzie pierwszy, a **M r a c z e w s k i** ostatni.
- cytaty
 - (8) Que votre nom soit béni à jamais, bien qui avez voulu m'éprouver par cette peine.



Nierozpoznane, charakterystyka: formy dawne

Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- fleksyjne:

Nierozpoznane, charakterystyka: formy dawne



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- fleksyjne:

(9) Rozumie się z doliczeniem procentu
w stosunku...sześć...do ośmiu o d s t a rocznie...



Nierozpoznane, charakterystyka: formy dawne

Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- fleksyjne:

- (9) Rozumie się z doliczeniem procentu w stosunku... sześć... do ośmiu o d s t a rocznie...
- (10) Pod oknem stał ten sam czarny stół obity suknem, także niegdyś z i e l o n e m, dziś tylko p o p l a m i o n e m.

Nierozpoznane, charakterystyka: formy dawne



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- fleksyjne:

- (9) Rozumie się z doliczeniem procentu w stosunku... sześć... do ośmiu o d s t a rocznie...
- (10) Pod oknem stał ten sam czarny stół obity suknem, także niegdyś z i e l o n e m, dziś tylko p o p l a m i o n e m.
- (11) Widzisz, mówię ci nawet k o m p l i m e n t a.

Nierozpoznane, charakterystyka: formy dawne



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

• fleksyjne:

- (9) Rozumie się z doliczeniem procentu w stosunku... sześć... do ośmiu o d s t a rocznie...
- (10) Pod oknem stał ten sam czarny stół obity suknem, także niegdyś z i e l o n e m, dziś tylko p o p l a m i o n e m.
- (11) Widzisz, mówię ci nawet k o m p l i m e n t a.
- (12) Zdublujesz majątek i jak mówi Stary Testament, zobaczysz n i e p r z y j a c i o t y twoje u podnóżka nóg twoich...

Nierozpoznane, charakterystyka: formy dawne



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

• ortograficzne

- (13) P r z y s z l ę ci tu zaraz Rzeckiego — odezwał się Wokulski, ściskając go za rękę.
- (14) ...ćwierćwiekowa pajęczyna, a z p e w n o ś c i ą ćwierćwiekowa firanka, niegdyś zielona...
- (15) W o g ó l e żaden na świecie język nie jest mi drogim i nie ma w moich oczach żadnych praw.
- (16) Mów, co chcesz, a ja wiem, że dałbym kilka lat życia, aby się d o w i e d z i é ć szczegółów o t é j kobiecie, aby m ó d z ją poznać.
- (17) Przez tę p r o w i n c y ą g ł ó w n a a r m i j a r o s s y j s k a miała swoje k o m m u n i k a c y e z Kijowem.



Rozszerzenie współczesnego analizatora o jednostki dawne

Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- korpus o długości 1 mln segmentów (1000 próbek gronowych po 1000 segmentów)
- analiza zgromadzonego materiału za pomocą niemodyfikowanego analizatora współczesnego
- stworzenie listy jednostek nierozpoznanych stanowiących potencjalnie formy dawne
- wzbogacenie zbioru form o hasła *Słownika warszawskiego*, nienotowane we współczesnych słownikach języka polskiego

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

**Podsumowanie:
synchronia i diachronia**

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich
- notowanie zmian dotyczących charakterystyki leksemów, np. zmian paradygmatu odmiany

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich
- notowanie zmian dotyczących charakterystyki leksemów, np. zmian paradygmatu odmiany
- oznaczenie form notowanych w tekstach dawnych

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:

synchronia i diachronia

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich
- notowanie zmian dotyczących charakterystyki leksemów, np. zmian paradygmatu odmiany
- oznaczenie form notowanych w tekstach dawnych
- oznaczenie jednostek dawnych (na materiale *Słownika wileńskiego*)

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:

synchronia i diachronia

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich
- notowanie zmian dotyczących charakterystyki leksemów, np. zmian paradygmatu odmiany
- oznaczenie form notowanych w tekstach dawnych
- oznaczenie jednostek dawnych (na materiale *Słownika wileńskiego*)

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich
- notowanie zmian dotyczących charakterystyki leksemów, np. zmian paradygmatu odmiany
- oznaczenie form notowanych w tekstach dawnych
- oznaczenie jednostek dawnych (na materiale *Słownika wileńskiego*)

(18) Szpanowali hipsterskiémi iPhony

Rozszerzenie współczesnego analizatora o jednostki dawne, cd.



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- przyporządkowanie jednostkom wzorców odmiany, a w przypadku braku wzorców — stworzenie ich
- notowanie zmian dotyczących charakterystyki leksemów, np. zmian paradygmatu odmiany
- oznaczenie form notowanych w tekstach dawnych
- oznaczenie jednostek dawnych (na materiale *Słownika wileńskiego*)

(18) Szpanowali hipsterskiémi iPhony

(19) W kyosku kupić ulubiony komix

Podsumowanie: synchronia i diachronia



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

- analizator morfologiczny uzupełniony o wzorce z 2 poł. XIX wieku
- informacja o ewolucji form
- model diachronicznego słownika fleksyjnego
- minikorpus



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia

Dziękujemy za uwagę!



Literatura I



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia



ACEDAŃSKI S., *A Morphosyntactic Brill Tagger for Inflectional Languages*, [w:] *Advances in Natural Language Processing. 7th International Conference on NLP, IceTAL 2010, Reykjavik, Iceland, August 16–18, 2010*, red. H. LOFTSSON, E. RÖGNVALDSSON i S. HELGADÓTTIR, s. 3–14, Springer .



JASSEM, K. *Przetwarzanie tekstów polskich w systemie tłumaczenia automatycznego POLENG*, Poznań 2012.



MAZIARZ M., PIASECKI M. i SZPAKOWICZ S., *Approaching plWordNet 2.0* 2012.



OBREŃBSKI T., *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*, Rozprawa doktorska, Instytut Podstaw Informatyki PAN 2002.



PIASECKI M., *Polish Tagger TaKIPI: Rule Based Construction and Optimisation*, „Task Quarterly”, t. 11 (1–2), s. 151–167 2007.



RABIEGA-WIŚNIEWSKA, J. i RUDOLF, M., *AMOR — program automatycznej analizy fleksyjnej tekstu polskiego*, „Biuletyn PTJ”, LXVIII, s. 175–186.



SGJP-II, SALONI Z., GRUSZCZYŃSKI W., WOLIŃSKI M., WOŁOSZ R. i MAKOWSKA D., *Słownik gramatyczny języka polskiego*, Warszawa 2012, cD.



SZAFRAN, K., *Analizator morfologiczny SAM-96, opis użytkowy*, Raport Instytutu Informatyki UW, TR 96-05(226), maj, 1996.



SJPDor, DOROSZEWSKI W. (red.), *Słownik języka polskiego*, t. I–XII, Warszawa 1958.

Literatura II



Zasób leksykalny XIX w.

M. Derwojedowa,
W. Kieraś,
D. Skowrońska,
R. Wołosz

Projekt

Narzędzia

Synchronia: zasoby
i narzędzia

Analizatory morfologiczne

Diachronia: zasoby i programy

Automatyczna analiza
morfologiczna

SJPDor

Materiał

Testy

Nierozpoznane

Podsumowanie:
synchronia i diachronia



SFJP, Lubaszewski, W. (red.) *Słownik fleksyjny języka polskiego*, Kraków 2001.



WOLIŃSKI M., *Komputerowa weryfikacja gramatyki Świdzińskiego*, Rozprawa doktorska, Instytut Podstaw Informatyki, Polska Akademia Nauk, Warsaw 2004.



WOLIŃSKI M., *Morfeusz – a Practical Tool for the Morphological Analysis of Polish*, [w:] *Intelligent Information Processing and Web Mining*, red. M.A. KŁOPOTEK, S.T. WIERZCHOŃ i K. TROJANOWSKI, *Advances in Soft Computing*, s. 503–512, Springer-Verlag, Berlin 2006.



WOLIŃSKI M., GŁOWIŃSKA K. i ŚWIDZIŃSKI M., *A Preliminary Version of Składnica – a Treebank of Polish*, [w:] *Proceedings of the 5th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, red. Z. VETULANI, s. 299–303, Poznań, Poland 2011.



Wołosz R., *Efektywna metoda analizy i syntezy morfologicznej w języku polskim*, 2005.