

Składniowe przetwarzanie tekstów polskich

Adam Przepiórkowski



INSTYTUT PODSTAW INFORMATYKI
POLSKIEJ AKADEMII NAUK
ul. Jana Kazimierza 5, 01-248 Warszawa

HLT Days 2012

28 września 2012



Prace w Zespole Inżynierii Lingwistycznej IPI PAN:

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

Hasłowo:

- znakowanie i przeszukiwanie korpusów, Narodowy Korpus Języka Polskiego, korpusy znakowane składniowo,
- zasoby leksykalne: słownik morfoskładniowy (PoliMorf), słowniki nazw własnych, słowniki walencyjne itp.,
- tagery, identyfikacja nazw własnych, parsery,
- identyfikacja nawiązań (ang. *coreference resolution*),
- ujednoznacznianie sensów słów, automatyczne streszczanie, ekstrakcja informacji, kategoryzacja tekstów,
- analiza wydźwięku (ang. *sentiment analysis*).



Prace w Zespole Inżynierii Lingwistycznej IPI PAN:

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

Hasłowo:

- znakowanie i przeszukiwanie korpusów, Narodowy Korpus Języka Polskiego, korpusy znakowane składniowo,
- zasoby leksykalne: słownik morfoskładniowy (PoliMorf), słowniki nazw własnych, słowniki walencyjne itp.,
- tagery, identyfikacja nazw własnych, **parsery**
- identyfikacja nawiązań (ang. *coreference resolution*),
- ujednoznacznianie sensów słów, automatyczne streszczanie, ekstrakcja informacji, kategoryzacja tekstów,
- analiza wydźwięku (ang. *sentiment analysis*).



Prace w Zespole Inżynierii Lingwistycznej IPI PAN:

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

Hasłowo:

- znakowanie i przeszukiwanie korpusów, Narodowy Korpus Języka Polskiego, korpusy znakowane składniowo,
- zasoby leksykalne: słownik morfoskładniowy (PoliMorf), słowniki nazw własnych, słowniki walencyjne itp.,
- tagery, identyfikacja nazw własnych, **parsery**,
- identyfikacja nawiązań (ang. *coreference resolution*),
- ujednoznacznianie sensów słów, automatyczne streszczanie, ekstrakcja informacji, kategoryzacja tekstów,
- analiza wydźwięku (ang. *sentiment analysis*).

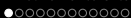


Przetwarzanie składniowe – uwagi historyczne



Długa historia, kilka (często zamkniętych) parserów:

- Stanisław Szpakowicz, 1978, *Automatyczna analiza składniowa zdań pisanych*, 1986, *Formalny opis składniowy zdań polskich*,
- Marek Świdziński, 1992, *Gramatyka formalna języka polskiego*, pierwsze implementacje na MIM UW w ramach projektów Janusza S. Bienia,
- równolegle, od lat 1980. – prace Zygmunta Vetulaniego na Uniwersytecie im. Adama Mickiewicza,
- Tomasz Obrębski, 2002, *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*,
- lata 2000. – prace Filipa Gralińskiego i Krzysztofa Jassem w firmie POLENG,
- lata 2000. – Nina Suszczańska, Przemysław Szmal – system Thetos.



Przetwarzanie składniowe – uwagi historyczne



Długa historia, kilka (często zamkniętych) parserów:

- Stanisław Szpakowicz, 1978, *Automatyczna analiza składniowa zdań pisanych*, 1986, *Formalny opis składniowy zdań polskich*,
- Marek Świdziński, 1992, *Gramatyka formalna języka polskiego*, pierwsze implementacje na MIM UW w ramach projektów Janusza S. Bienia,
- równolegle, od lat 1980. – prace Zygmunta Vetulaniego na Uniwersytecie im. Adama Mickiewicza,
- Tomasz Obrębcki, 2002, *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*,
- lata 2000. – prace Filipa Gralińskiego i Krzysztofa Jassemę w firmie POLENG,
- lata 2000. – Nina Suszczańska, Przemysław Szmal – system Thetos.

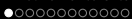


Przetwarzanie składniowe – uwagi historyczne



Długa historia, kilka (często zamkniętych) parserów:

- Stanisław Szpakowicz, 1978, *Automatyczna analiza składniowa zdań pisanych*, 1986, *Formalny opis składniowy zdań polskich*,
- Marek Świdziński, 1992, *Gramatyka formalna języka polskiego*, pierwsze implementacje na MIM UW w ramach projektów Janusza S. Bienia,
- równolegle, od lat 1980. – prace Zygmunta Vetulaniego na Uniwersytecie im. Adama Mickiewicza,
- Tomasz Obrębski, 2002, *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*,
- lata 2000. – prace Filipa Galińskiego i Krzysztofa Jassem w firmie POLENG,
- lata 2000. – Nina Suszczańska, Przemysław Szmalec – system Thetos.

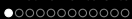


Przetwarzanie składniowe – uwagi historyczne



Długa historia, kilka (często zamkniętych) parserów:

- Stanisław Szpakowicz, 1978, *Automatyczna analiza składniowa zdań pisanych*, 1986, *Formalny opis składniowy zdań polskich*,
- Marek Świdziński, 1992, *Gramatyka formalna języka polskiego*, pierwsze implementacje na MIM UW w ramach projektów Janusza S. Bienia,
- równolegle, od lat 1980. – prace Zygmunta Vetulaniego na Uniwersytecie im. Adama Mickiewicza,
- Tomasz Obrębski, 2002, *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*,
- lata 2000. – prace Filipa Gralińskiego i Krzysztofa Jassem w firmie POLENG,
- lata 2000. – Nina Suszczańska, Przemysław Szmaj – system Thetos.



Przetwarzanie składniowe – uwagi historyczne



Długa historia, kilka (często zamkniętych) parserów:

- Stanisław Szpakowicz, 1978, *Automatyczna analiza składniowa zdań pisanych*, 1986, *Formalny opis składniowy zdań polskich*,
- Marek Świdziński, 1992, *Gramatyka formalna języka polskiego*, pierwsze implementacje na MIM UW w ramach projektów Janusza S. Bienia,
- równolegle, od lat 1980. – prace Zygmunta Vetulaniego na Uniwersytecie im. Adama Mickiewicza,
- Tomasz Obrębski, 2002, *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*,
- lata 2000. – prace Filipa Gralińskiego i Krzysztofa Jassem w firmie POLENG,
- lata 2000. – Nina Suszczańska, Przemysław Szmalec – system Thetos.

Przetwarzanie składniowe – uwagi historyczne



Długa historia, kilka (często zamkniętych) parserów:

- Stanisław Szpakowicz, 1978, *Automatyczna analiza składniowa zdań pisanych*, 1986, *Formalny opis składniowy zdań polskich*,
- Marek Świdziński, 1992, *Gramatyka formalna języka polskiego*, pierwsze implementacje na MIM UW w ramach projektów Janusza S. Bienia,
- równolegle, od lat 1980. – prace Zygmunta Vetulaniego na Uniwersytecie im. Adama Mickiewicza,
- Tomasz Obrębski, 2002, *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*,
- lata 2000. – prace Filipa Gralińskiego i Krzysztofa Jassemę w firmie POLENG,
- lata 2000. – Nina Suszczańska, Przemysław Szmalec – system Thetos.



Przetwarzanie składniowe w IPI PAN



1. HPSG

Gramatyka w formalizmie *Head-driven Phrase Structure Grammar* (Pollard and Sag, 1987, 1994), rozwijana w ZIL IPI PAN od pierwszej połowy lat 1990.:

- Adam Przepiórkowski, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka, 2002, *Formalny opis języka polskiego: Teoria i implementacja*,
- bardzo bogaty opis lingwistyczny,
- doktoraty: A. Mykowiecka, A. Przepiórkowski, A. Kupść, M. Marciniak,
- implementacja typu *toy grammar*.

Intensywne prace do początku lat 2000.



Przetwarzanie składniowe w IPI PAN



1. HPSG

Gramatyka w formalizmie *Head-driven Phrase Structure Grammar* (Pollard and Sag, 1987, 1994), rozwijana w ZIL IPI PAN od pierwszej połowy lat 1990.:

- Adam Przepiórkowski, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka, 2002, *Formalny opis języka polskiego: Teoria i implementacja*,
- bardzo bogaty opis lingwistyczny
- doktoraty: A. Mykowiecka, A. Przepiórkowski, A. Kupść, M. Marciniak,
- implementacja typu *toy grammar*.

Intensywne prace do początku lat 2000.

Przetwarzanie składniowe w IPI PAN



1. HPSG

Gramatyka w formalizmie *Head-driven Phrase Structure Grammar* (Pollard and Sag, 1987, 1994), rozwijana w ZIL IPI PAN od pierwszej połowy lat 1990.:

- Adam Przepiórkowski, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka, 2002, *Formalny opis języka polskiego: Teoria i implementacja*,
- bardzo bogaty opis lingwistyczny,
- doktoraty: A. Mykowiecka, A. Przepiórkowski, A. Kupść, M. Marciniak,
- implementacja typu *toy grammar*.

Intensywne prace do początku lat 2000.



1. HPSG

Gramatyka w formalizmie *Head-driven Phrase Structure Grammar* (Pollard and Sag, 1987, 1994), rozwijana w ZIL IPI PAN od pierwszej połowy lat 1990.:

- Adam Przepiórkowski, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka, 2002, *Formalny opis języka polskiego: Teoria i implementacja*,
- bardzo bogaty opis lingwistyczny,
- doktoraty: A. Mykowiecka, A. Przepiórkowski, A. Kupść, M. Marciniak,
- implementacja typu *toy grammar*.

Intensywne prace do początku lat 2000.



1. HPSG

Gramatyka w formalizmie *Head-driven Phrase Structure Grammar* (Pollard and Sag, 1987, 1994), rozwijana w ZIL IPI PAN od pierwszej połowy lat 1990.:

- Adam Przepiórkowski, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka, 2002, *Formalny opis języka polskiego: Teoria i implementacja*,
- bardzo bogaty opis lingwistyczny,
- doktoraty: A. Mykowiecka, A. Przepiórkowski, A. Kupść, M. Marciniak,
- implementacja typu *toy grammar*.

Intensywne prace do początku lat 2000.



2. Świga

Implementacja gramatyki Marka Świdzińskiego:

- Marcin Woliński, 2004, *Komputerowa weryfikacja gramatyki Świdzińskiego*, Świga,
- Maciej Ogrodniczuk, 2006, *Weryfikacja korpusu wypowiedników polskich (z wykorzystaniem gramatyki formalnej Świdzińskiego)*, Uniwersytet Warszawski,
- gramatyka dalej rozwijana, z udziałem prof. Świdzińskiego, w ramach (pod)projektów M. Wolińskiego,
- rozwój skorelowany z budową banku drzew, *Składnica*,
- wyszukiwarka drzew składniowych.



2. Świga

Implementacja gramatyki Marka Świdzińskiego:

- Marcin Woliński, 2004, *Komputerowa weryfikacja gramatyki Świdzińskiego*, Świga,
- Maciej Ogrodniczuk, 2006, *Weryfikacja korpusu wypowiedników polskich (z wykorzystaniem gramatyki formalnej Świdzińskiego)*, Uniwersytet Warszawski,
- gramatyka dalej rozwijana, z udziałem prof. Świdzińskiego, w ramach (pod)projektów M. Wolińskiego,
- rozwój skorelowany z budową banku drzew, *Składnica*,
- wyszukiwarka drzew składniowych.



2. Świga

Implementacja gramatyki Marka Świdzińskiego:

- Marcin Woliński, 2004, *Komputerowa weryfikacja gramatyki Świdzińskiego*, Świga,
- Maciej Ogrodniczuk, 2006, *Weryfikacja korpusu wypowiedników polskich (z wykorzystaniem gramatyki formalnej Świdzińskiego)*, Uniwersytet Warszawski,
- gramatyka dalej rozwijana, z udziałem prof. Świdzińskiego, w ramach (pod)projektów M. Wolińskiego,
- rozwój skorelowany z budową banku drzew, *Składnica*,
- wyszukiwarka drzew składniowych.

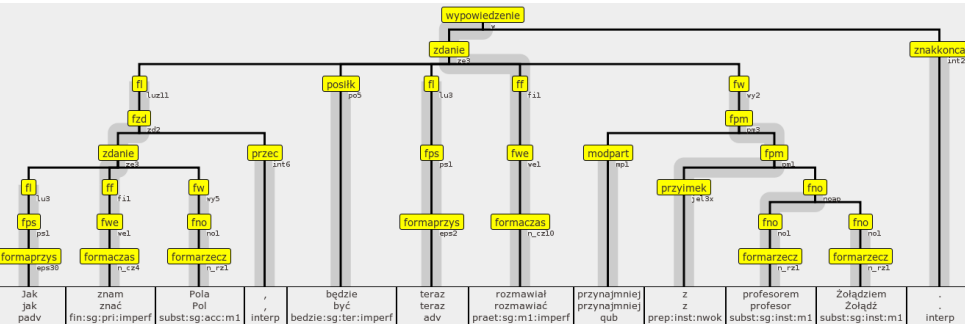


2. Świga

Implementacja gramatyki Marka Świdzińskiego:

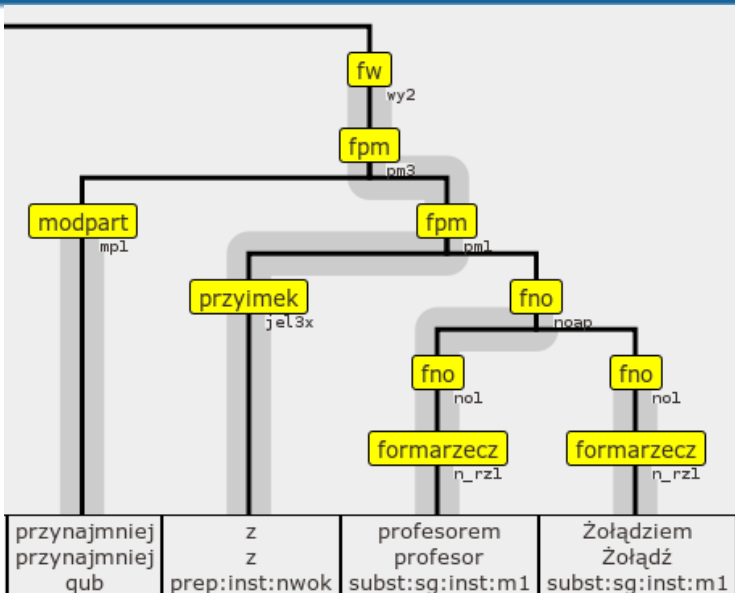
- Marcin Woliński, 2004, *Komputerowa weryfikacja gramatyki Świdzińskiego*, Świga,
- Maciej Ogrodniczuk, 2006, *Weryfikacja korpusu wypowiedników polskich (z wykorzystaniem gramatyki formalnej Świdzińskiego)*, Uniwersytet Warszawski,
- gramatyka dalej rozwijana, z udziałem prof. Świdzińskiego, w ramach (pod)projektów M. Wolińskiego,
- rozwój skorelowany z budową banku drzew, *Składnica*,
- wyszukiwarka drzew składniowych.

Przetwarzanie składniowe w IPI PAN



Jak znam Pola, będzie teraz rozmawiał przynajmniej z profesorem Złodźciem.

Przetwarzanie składniowe w IPI PAN





3. Przetwarzanie zależnościowe

Prace Aliny Wróblewskiej – 2 metody (pół)automatycznego tworzenia zależnościowego banku drzew:

- konwersja składnikowego banku drzew (A. Wróblewska, M. Woliński, SIIS 2011),
- wykorzystanie korpusów równoległych i narzędzi dla języka angielskiego.



3. Przetwarzanie zależnościowe

Prace Aliny Wróblewskiej – 2 metody (pół)automatycznego tworzenia zależnościowego banku drzew:

- konwersja składnikowego banku drzew (A. Wróblewska, M. Woliński, SIIS 2011),
- wykorzystanie korpusów równoległych i narzędzi dla języka angielskiego.



3. Przetwarzanie zależnościowe

Prace Aliny Wróblewskiej – 2 metody (pół)automatycznego tworzenia zależnościowego banku drzew:

- konwersja składnikowego banku drzew (A. Wróblewska, M. Woliński, SIIS 2011),
- wykorzystanie korpusów równoległych i narzędzi dla języka angielskiego.

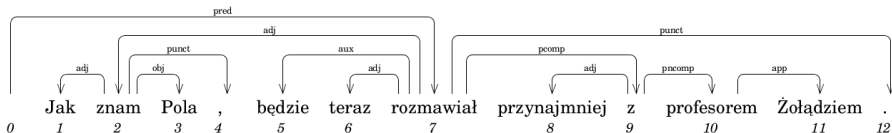
Przetwarzanie składniowe w IPI PAN



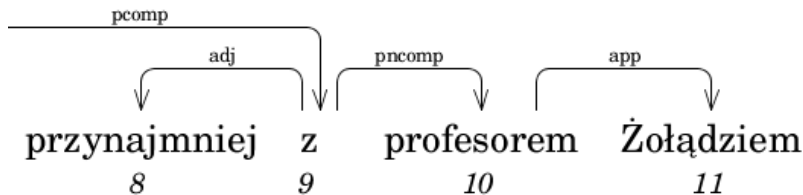
3. Przetwarzanie zależnościowe

Prace Aliny Wróblewskiej – 2 metody (pół)automatycznego tworzenia zależnościowego banku drzew:

- konwersja składnikowego banku drzew (A. Wróblewska, M. Woliński, SIIS 2011),
- wykorzystanie korpusów równoległych i narzędzi dla języka angielskiego.



Przetwarzanie składniowe w IPI PAN



Przetwarzanie składniowe w IPI PAN



Wykorzystanie korpusów równoległych, np.:

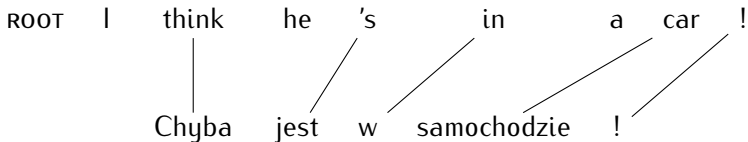
- *I think he's in a car!*
- *Chyba jest w samochodzie!*

Przetwarzanie składniowe w IPI PAN



Wykorzystanie korpusów równoległych, np.:

- *I think he's in a car!*
- *Chyba jest w samochodzie!*

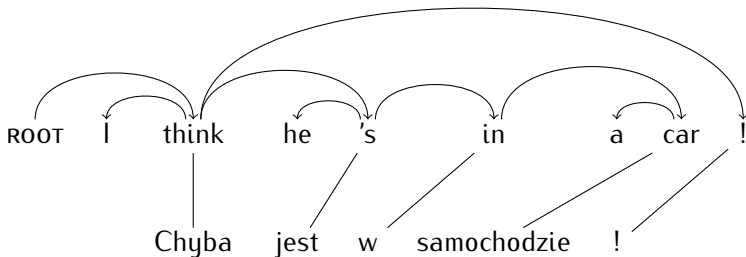


Przetwarzanie składniowe w IPI PAN



Wykorzystanie korpusów równoległych, np.:

- *I think he's in a car!*
- *Chyba jest w samochodzie!*

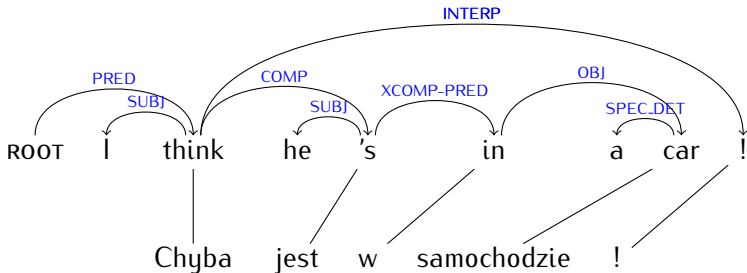


Przetwarzanie składniowe w IPI PAN



Wykorzystanie korpusów równoległych, np.:

- *I think he's in a car!*
- *Chyba jest w samochodzie!*

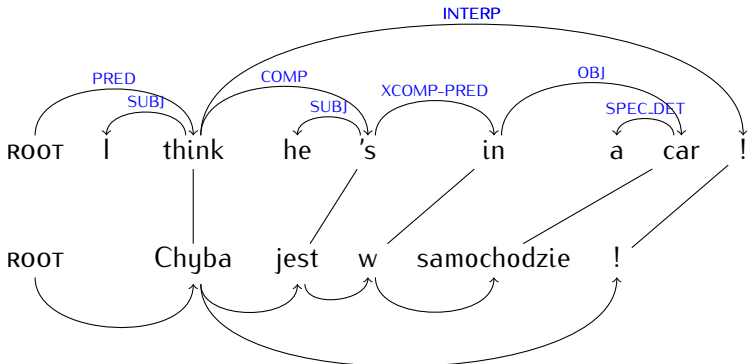


Przetwarzanie składniowe w IPI PAN



Wykorzystanie korpusów równoległych, np.:

- *I think he's in a car!*
- *Chyba jest w samochodzie!*

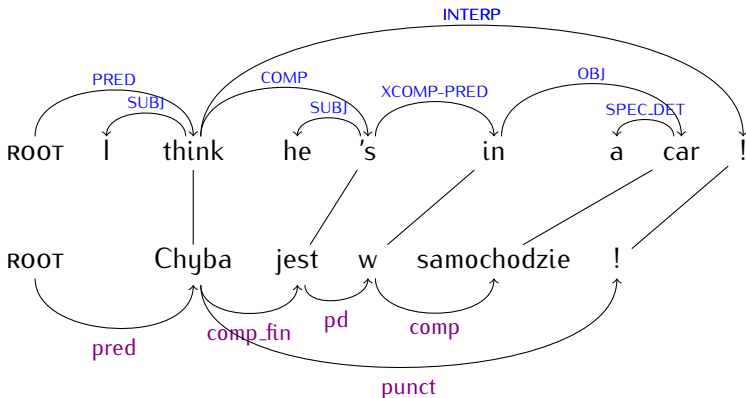


Przetwarzanie składniowe w IPI PAN



Wykorzystanie korpusów równoległych, np.:

- *I think he's in a car!*
- *Chyba jest w samochodzie!*





4. LFG

Prace Agnieszki Patejuk i autora – w Lexical-Functional Grammar (Bresnan, 1982; Dalrymple, 2001):

- 2 główne poziomy składniowe:
 - składnikowy (*c-structure*),
 - funkcyjny (*f-structure*),
- hasła słownikowe i reguły składniowe definiują oba poziomy równoległe.



4. LFG

Prace Agnieszki Patejuk i autora – w Lexical-Functional Grammar (Bresnan, 1982; Dalrymple, 2001):

- 2 główne poziomy składniowe:
 - składnikowy (*c-structure*),
 - funkcyjny (*f-structure*),
- hasła słownikowe i reguły składniowe definiują oba poziomy równolegle.



4. LFG

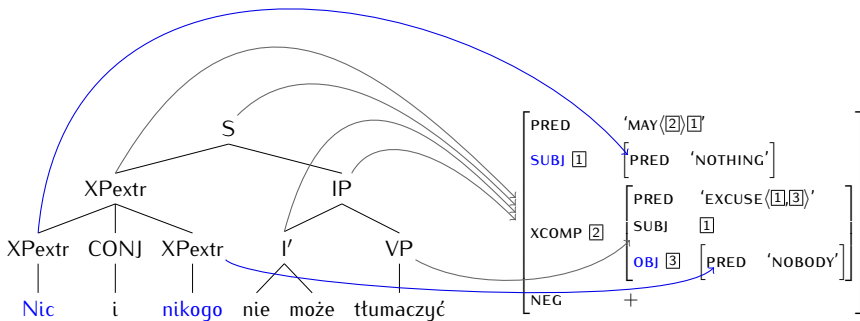
Prace Agnieszki Patejuk i autora – w Lexical-Functional Grammar (Bresnan, 1982; Dalrymple, 2001):

- 2 główne poziomy składniowe:
 - składnikowy (*c-structure*),
 - funkcyjny (*f-structure*),
- hasła słownikowe i reguły składniowe definiują oba poziomy równoległe.

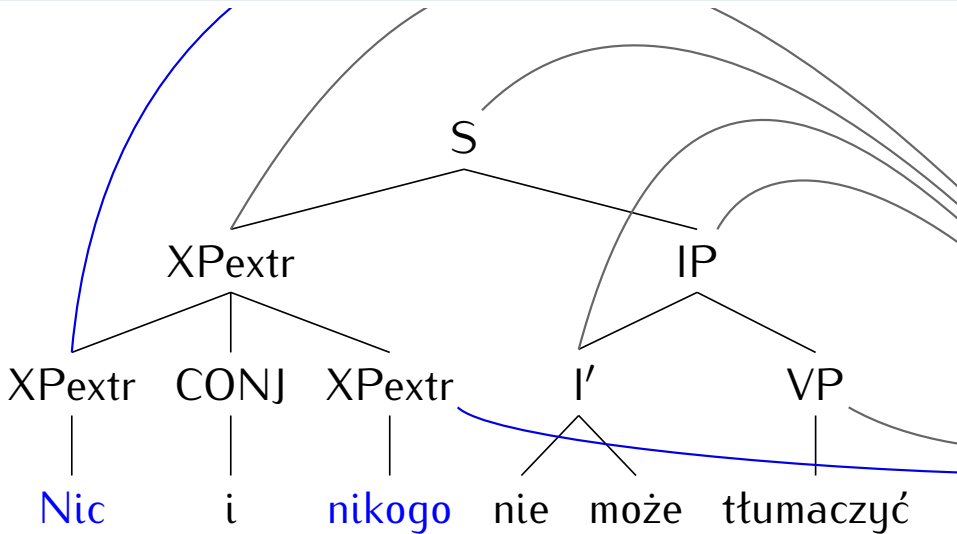
Przetwarzanie składniowe w IPI PAN



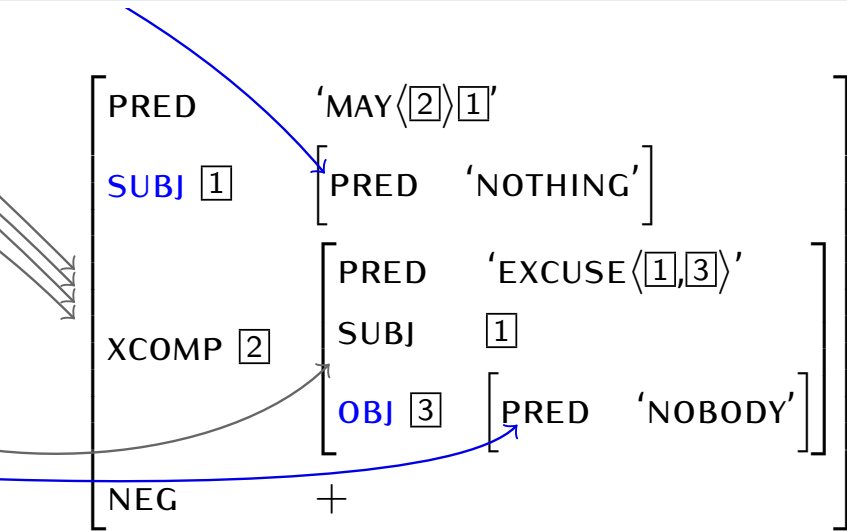
$XP_{extr}^{neg} \rightarrow XP_{extr}^{neg} \text{ CONJ } XP_{extr}^{neg}$
 $\uparrow=\downarrow \qquad \qquad \qquad \uparrow=\downarrow$



Przetwarzanie składniowe w IPI PAN



Przetwarzanie składniowe w IPI PAN



Przetwarzanie składniowe w IPI PAN



Pasożytnicze podejście do rozwoju gramatyki:

- obecne wersje *Świgrzy* – pierwsze wersje struktury składnikowej,
- gramatyka HPSG – pierwsze wersje struktury funkcyjnej,
- Morfeusz SGJP (Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz) – źródło informacji morfosyntaktycznych,
- NKJP i Składnica – testowanie, znajdowanie przykładów itd.; także źródło informacji morfoskładniowych.

Przetwarzanie składniowe w IPI PAN



Pasożytnicze podejście do rozwoju gramatyki:

- obecne wersje *Świgrzy* – pierwsze wersje struktury składnikowej,
- gramatyka HPSG – pierwsze wersje struktury funkcyjnej,
- Morfeusz SGJP (Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz) – źródło informacji morfosyntaktycznych,
- NKJP i Składnica – testowanie, znajdowanie przykładów itd.; także źródło informacji morfoskładniowych.

Przetwarzanie składniowe w IPI PAN



Pasożytnicze podejście do rozwoju gramatyki:

- obecne wersje *Świgrzy* – pierwsze wersje struktury składnikowej,
- gramatyka HPSG – pierwsze wersje struktury funkcyjnej,
- Morfeusz SGJP (Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz) – źródło informacji morfosyntaktycznych,
- NKJP i Składnica – testowanie, znajdowanie przykładów itd.; także źródło informacji morfoskładniowych.



Przetwarzanie składniowe w IPI PAN



Pasożytnicze podejście do rozwoju gramatyki:

- obecne wersje *Świgrzy* – pierwsze wersje struktury składnikowej,
- gramatyka HPSG – pierwsze wersje struktury funkcyjnej,
- Morfeusz SGJP (Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz) – źródło informacji morfosyntaktycznych,
- NKJP i Składnica – testowanie, znajdowanie przykładów itd.; także źródło informacji morfoskładniowych.



Przetwarzanie składniowe w IPI PAN



Pasożytnicze podejście do rozwoju gramatyki:

- obecne wersje *Świgrzy* – pierwsze wersje struktury składnikowej,
- gramatyka HPSG – pierwsze wersje struktury funkcyjnej,
- Morfeusz SGJP (Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz) – źródło informacji morfosyntaktycznych,
- NKJP i Składnica – testowanie, znajdowanie przykładów itd.; także źródło informacji morfoskładniowych.



Przetwarzanie składniowe w IPI PAN



Pasożytnicze podejście do rozwoju gramatyki:

- obecne wersje *Świgrzy* – pierwsze wersje struktury składnikowej,
- gramatyka HPSG – pierwsze wersje struktury funkcyjnej,
- Morfeusz SGJP (Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz) – źródło informacji morfosyntaktycznych,
- NKJP i Składnica – testowanie, znajdowanie przykładów itd.; także źródło informacji morfoskładniowych.



Pominąłem intensywne prace dotyczące:

- parsowania powierzchniowego (system Spejd),
- budowy słownika walencyjnego (Filip Skwarski i in.; zob. wczorajsze plakaty),
- automatycznego wykrywania błędów w bankach drzew (Katarzyna Krasnowska i in.; TSD 2012).

Przyszłość:

- parsery probabilistyczne (już teraz trenowane są parsery zależnościowe),
- semantyka.

Przetwarzanie składniowe w IPI PAN



Pomiąłem intensywne prace dotyczące:

- parsowania powierzchniowego (system Spejd),
- budowy słownika walencyjnego (Filip Skwarski i in.; zob. wczorajsze plakaty),
- automatycznego wykrywania błędów w bankach drzew (Katarzyna Krasnowska i in.; TSD 2012).

Przyszłość:

- parsery probabilistyczne (już teraz trenowane są parsery zależnościowe),
- semantyka.

Przetwarzanie składniowe w IPI PAN



Pominąłem intensywne prace dotyczące:

- parsowania powierzchniowego (system Spejd),
- budowy słownika walencyjnego (Filip Skwarski i in.; zob. wczorajsze plakaty),
- automatycznego wykrywania błędów w bankach drzew (Katarzyna Krasnowska i in.; TSD 2012).

Przyszłość:

- parsery probabilistyczne (już teraz trenowane są parsery zależnościowe),
- semantyka.



Przetwarzanie składniowe w IPI PAN



Pominąłem intensywne prace dotyczące:

- parsowania powierzchniowego (system Spejd),
- budowy słownika walencyjnego (Filip Skwarski i in.; zob. wczorajsze plakaty),
- automatycznego wykrywania błędów w bankach drzew (Katarzyna Krasnowska i in.; TSD 2012).

Przyszłość:

- parsery probabilistyczne (już teraz trenowane są parsery zależnościowe),
- semantyka.



Przetwarzanie składniowe w IPI PAN

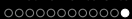


Pominąłem intensywne prace dotyczące:

- parsowania powierzchniowego (system Spejd),
- budowy słownika walencyjnego (Filip Skwarski i in.; zob. wczorajsze plakaty),
- automatycznego wykrywania błędów w bankach drzew (Katarzyna Krasnowska i in.; TSD 2012).

Przyszłość:

- parsery probabilistyczne (już teraz trenowane są parsery zależnościowe),
- semantyka.



Przetwarzanie składniowe w IPI PAN



Pominąłem intensywne prace dotyczące:

- parsowania powierzchniowego (system Spejd),
- budowy słownika walencyjnego (Filip Skwarski i in.; zob. wczorajsze plakaty),
- automatycznego wykrywania błędów w bankach drzew (Katarzyna Krasnowska i in.; TSD 2012).

Przyszłość:

- parsery probabilistyczne (już teraz trenowane są parsery zależnościowe),
- semantyka.

Zamiast podsumowania



- Zespół Inżynierii Lingwistycznej IPI PAN:
<http://zil.ipipan.waw.pl/>,
- Computational Linguistics in Poland:
<http://clip.ipipan.waw.pl/>,
- Journal of Language Modelling:
<http://jlm.ipipan.waw.pl/>,
- Narodowy Korpus Języka Polskiego:
<http://nkjp.pl/>,
- Polska Lista Językoznawcza:
<http://tnij.org/pling>.

Zamiast podsumowania



- Zespół Inżynierii Lingwistycznej IPI PAN:
<http://zil.ipipan.waw.pl/>,
- Computational Linguistics in Poland:
<http://clip.ipipan.waw.pl/>,
- Journal of Language Modelling:
<http://jlm.ipipan.waw.pl/>,
- Narodowy Korpus Języka Polskiego:
<http://nkjp.pl/>,
- Polska Lista Językoznawcza:
<http://tnij.org/pling>.

Zamiast podsumowania



- Zespół Inżynierii Lingwistycznej IPI PAN:
<http://zil.ipipan.waw.pl/>,
- Computational Linguistics in Poland:
<http://clip.ipipan.waw.pl/>,
- Journal of Language Modelling:
<http://jlm.ipipan.waw.pl/>,
- Narodowy Korpus Języka Polskiego:
<http://nkjp.pl/>,
- Polska Lista Językoznawcza:
<http://tnij.org/pling>.

Zamiast podsumowania



- Zespół Inżynierii Lingwistycznej IPI PAN:
<http://zil.ipipan.waw.pl/>,
- Computational Linguistics in Poland:
<http://clip.ipipan.waw.pl/>,
- Journal of Language Modelling:
<http://jlm.ipipan.waw.pl/>,
- Narodowy Korpus Języka Polskiego:
<http://nkjp.pl/>,
- Polska Lista Językoznawcza:
<http://tnij.org/pling>.

Zamiast podsumowania



- Zespół Inżynierii Lingwistycznej IPI PAN:
<http://zil.ipipan.waw.pl/>,
- Computational Linguistics in Poland:
<http://clip.ipipan.waw.pl/>,
- Journal of Language Modelling:
<http://jlm.ipipan.waw.pl/>,
- Narodowy Korpus Języka Polskiego:
<http://nkjp.pl/>,
- Polska Lista Językoznawcza:
<http://tnij.org/pling>.

Zamiast podsumowania



- Zespół Inżynierii Lingwistycznej IPI PAN:
<http://zil.ipipan.waw.pl/>,
- Computational Linguistics in Poland:
<http://clip.ipipan.waw.pl/>,
- Journal of Language Modelling:
<http://jlm.ipipan.waw.pl/>,
- Narodowy Korpus Języka Polskiego:
<http://nkjp.pl/>,
- Polska Lista Językoznawcza:
<http://tnij.org/pling>.

Dziękuję za uwagę!

- Bień, J. S. (1997). Komputerowa weryfikacja formalnej gramatyki Świdzińskiego. *Biuletyn Polskiego Towarzystwa Językoznawczego*, LII, 147–164.
- Bresnan, J., editor (1982). *The Mental Representation of Grammatical Relations*. MIT Press Series on Cognitive Theory and Mental Representation. The MIT Press, Cambridge, MA.
- Dalrymple, M. (2001). *Lexical-Functional Grammar*. Academic Press.
- Graliński, F. (2002). Wstępujący parser języka polskiego na potrzeby systemu POLENG. In G. Demenko, M. Karpiński, and K. Jassem, editors, *Speech and Language Technology. Volume 6*. Polskie Towarzystwo Fonetyczne, Poznań.
- Graliński, F. (2007). *Formalizacja nieciągłości zdań przy zastosowaniu rozszerzonej gramatyki bezkontekstowej*. Rozprawa doktorska, Uniwersytet im. Adama Mickiewicza, Poznań.
- Jassem, K. (2006). *Przetwarzanie tekstów polskich w systemie tłumaczenia automatycznego POLENG*. Wydawnictwo Naukowe UAM, Poznań.
- Kupść, A. (2000). *An HPSG Grammar of Polish Clitics*. Rozprawa doktorska, Polska Akademia Nauk i Université Paris 7.
- Marciniak, M. (2001). *Algorytmy implementacyjne syntaktycznych reguł koreferencji zaimków dla języka polskiego w terminach HPSG*. Rozprawa doktorska, Instytut Podstaw Informatyki PAN.
- Mykowiecka, A. (1999). *Opis składniowy polskich konstrukcji względnych w formalizmie HPSG*. Rozprawa doktorska, Instytut Podstaw Informatyki PAN.

- Obrębski, T. (2002). *Automatyczna analiza składniowa języka polskiego z wykorzystaniem gramatyki zależnościowej*. Rozprawa doktorska, Instytut Podstaw Informatyki PAN, Warszawa.
- Ogrodniczuk, M. (2006). *Weryfikacja korpusu wypowiedników polskich (z wykorzystaniem gramatyki formalnej Świdzińskiego)*. Rozprawa doktorska, Uniwersytet Warszawski, Warszawa.
- Pollard, C. and Sag, I. A. (1987). *Information-Based Syntax and Semantics, Volume 1: Fundamentals*. Number 13 in CSLI Lecture Notes. CSLI Publications, Stanford, CA.
- Pollard, C. and Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. Chicago University Press / CSLI Publications, Chicago, IL.
- Przepiórkowski, A. (1999). *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Rozprawa doktorska, Universität Tübingen.
- Przepiórkowski, A., Kupść, A., Marciniak, M., and Mykowiecka, A. (2002). *Formalny opis języka polskiego: Teoria i implementacja*. Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Saloni, Z., Gruszczyński, W., Woliński, M., and Wołosz, R. (2007). *Słownik gramatyczny języka polskiego*. Wiedza Powszechna, Warszawa.
- Suszczańska, N. and Szmaj, P. (2005). Categorical grammar elements in the Thetos system's parser. In Z. Vetulani, editor, *Proceedings of the 2nd Language & Technology Conference*, pages 338–342, Poznań, Poland.

- Szapkowicz, S. (1978). *Automatyczna analiza składniowa zdań pisanych*. Rozprawa doktorska, Uniwersytet Warszawski, Warszawa.
- Szapkowicz, S. (1986). *Formalny opis składniowy zdań polskich*. Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Świdziński, M. (1992). *Gramatyka formalna języka polskiego*, volume 349 of *Rozprawy Uniwersytetu Warszawskiego*. Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Świdziński, M. and Woliński, M. (2010). Towards a bank of constituent parse trees for Polish. In P. Sojka, A. Horák, I. Kopeček, and K. Pala, editors, *Text, Speech and Dialogue: 13th International Conference, TSD 2010, Brno, Czech Republic*, Lecture Notes in Artificial Intelligence, pages 197–204, Berlin. Springer-Verlag.
- Vetulani, Z. (2004). *Komunikacja człowieka z maszyną. Komputerowe modelowanie kompetencji językowej*. Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Woliński, M. (2004). *Komputerowa weryfikacja gramatyki Świdzińskiego*. Rozprawa doktorska, Instytut Podstaw Informatyki PAN, Warszawa.
- Woliński, M. (2006). Morfeusz — a practical tool for the morphological analysis of Polish. In M. A. Kłopotek, S. T. Wierzchoń, and K. Trojanowski, editors, *Intelligent Information Processing and Web Mining*, Advances in Soft Computing, pages 503–512. Springer-Verlag, Berlin.

- Woliński, M. (2010). Dendrarium – an open source tool for treebank building. In M. A. Kłopotek, M. Marciniak, A. Mykowiecka, W. Penczek, and S. T. Wierchoń, editors, *Proceedings of IIS'2010*, pages 193–204.
- Wróblewska, A. and Woliński, M. (2011). Preliminary experiments in Polish dependency parsing. In P. Bouvry, M. A. Kłopotek, F. Leprevost, M. Marciniak, A. Mykowiecka, and H. Rybiński, editors, *Security and Intelligent Information Systems: International Joint Conference, SIIS 2011, Warsaw, Poland, June 13-14, 2011, Revised Selected Papers*, volume 7053 of *Lecture Notes in Computer Science*, pages 279–292. Springer-Verlag.