

Lingwistyka teoretyczna w erze dużych modeli językowych

Adam Przepiórkowski



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



Obchody 50-lecia Instytutu Podstaw Informatyki PAN
Warszawa, 28 września 2026

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw **Informatyki** PAN?!

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!



- dyscypliny związane ze sobą genetycznie

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!



- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!

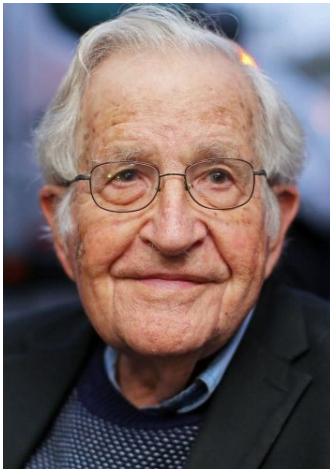


- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**
językoznawca, filozof, działacz polityczny

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!

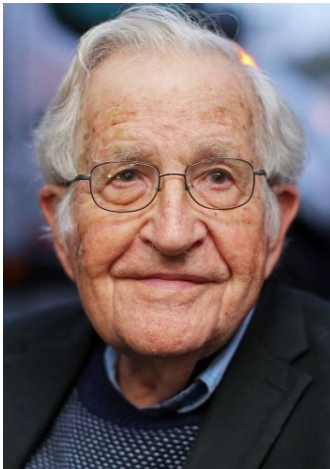


- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**
językoznawca, filozof, działacz polityczny
- **informatyka:**

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!

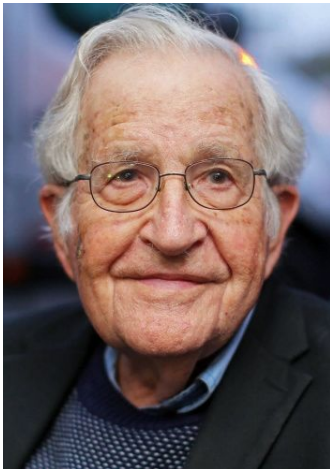


- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!



- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów
 - kluczowe znaczenie dla teorii języków programowania

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!

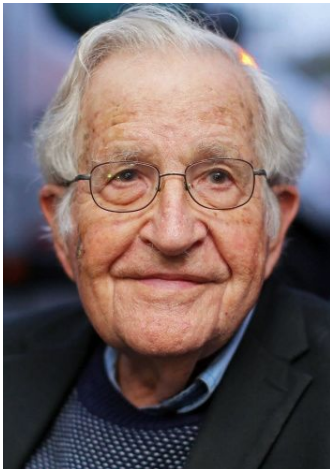


- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów
 - kluczowe znaczenie dla teorii języków programowania
- **lingwistyka:**

Lingwistyka w IPI PAN 1



Lingwistyka w Instytucie Podstaw Informatyki PAN?!



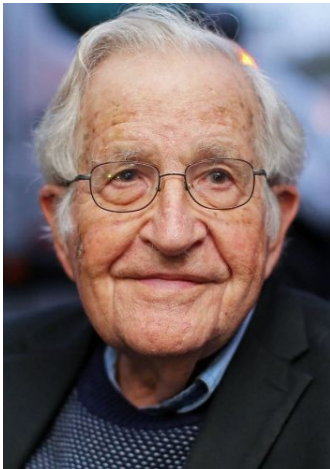
- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky (1928–)**
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów
 - kluczowe znaczenie dla teorii języków programowania
- **lingwistyka:**
 - ojciec lingwistyki formalnej, generatywnej,

Lingwistyka w IPI PAN 1

**Lingwistyka** w Instytucie Podstaw **Informatyki** PAN?!

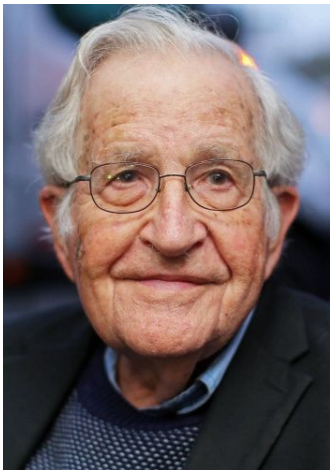
- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky** (1928–)
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów
 - kluczowe znaczenie dla teorii języków programowania
- **lingwistyka:**
 - ojciec lingwistyki formalnej, generatywnej,
 - teoretycznej?, współczesnej???, ...

Lingwistyka w IPI PAN 1

**Lingwistyka** w Instytucie Podstaw **Informatyki** PAN?!

- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky** (1928–)
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów
 - kluczowe znaczenie dla teorii języków programowania
- **lingwistyka:**
 - ojciec lingwistyki formalnej, generatywnej,
 - teoretycznej?, współczesnej???, ...
 - autor kolejnych wersji teorii składni „transformacyjnej”
(1957, 1965, 1981, 1986, 1995, 2001)

Lingwistyka w IPI PAN 1

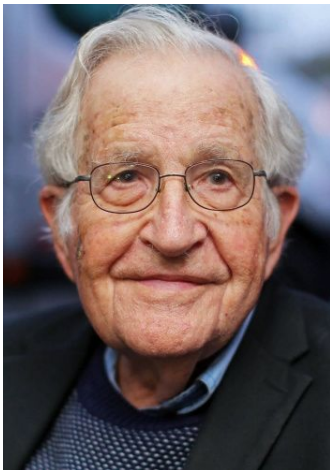
**Lingwistyka** w Instytucie Podstaw **Informatyki** PAN?!

- **dyscypliny związane ze sobą genetycznie**
- **Noam Chomsky** (1928–)
językoznawca, filozof, działacz polityczny
- **informatyka:**
 - hierarchia języków i odpowiadających im automatów
 - kluczowe znaczenie dla teorii języków programowania
- **lingwistyka:**
 - ojciec lingwistyki formalnej, generatywnej,
 - teoretycznej?, współczesnej???, ...
 - autor kolejnych wersji teorii składni „transformacyjnej” (1957, 1965, 1981, 1986, 1995, 2001)
 - w szeroko rozumianym językoznawstwie postać bardzo wpływowa, ale też bardzo kontrowersyjna

Lingwistyka w IPI PAN 1



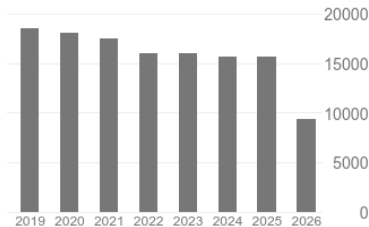
Lingwistyka w Instytucie Podstaw Informatyki PAN?!



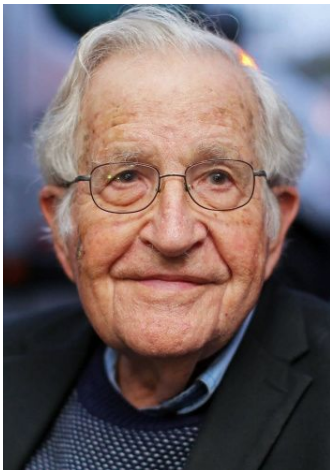
Cited by

[VIEW ALL](#)

	All	Since 2021
Citations	517002	90471
h-index	199	103
i10-index	1057	407



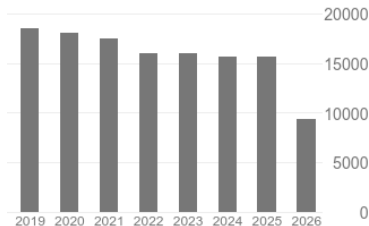
Lingwistyka w Instytucie Podstaw Informatyki PAN?!



Cited by

[VIEW ALL](#)

	All	Since 2021
Citations	517002	90471
h-index	199	103
i10-index	1057	407

[ABOUT](#) [CURRENT](#) [ALL ISSUES](#) [ANNOUNCEMENTS](#)[Home](#) / [Archives](#) / Vol. 9 No. 1 (2021) / [Articles](#)

Simplicity and the form of grammars

*Noam Chomsky***KEYWORDS:***simplicity, explanation, evaluation, grammar, faculty of language, learnability, evolvability, externalization***ABSTRACT**

The goal of theory construction is explanation: for language, theory for particular languages (grammar) and for the faculty of language FoL (the innate endowment for language acquisition). A primitive notion of simplicity of grammars is number of symbols, but this is too crude. An improved measure distinguishes grammars that capture genuine properties of language from those that do not. The theory of FoL must meet the empirical conditions of learnability (under extreme poverty of stimulus), and evolvability (given the limited but not insignificant evidence available). Recent work provides promising insights into how these twin conditions may be satisfied.

Lingwistyka w IPI PAN 2



Ale **skąd** lingwistyka w IPI PAN?

Lingwistyka w IPI PAN 2



Ale **skąd lingwistyka w IPI PAN?**

- Leonard Bolc (1934–2013)



Lingwistyka w IPI PAN 2



Ale **skąd lingwistyka w IPI PAN?**



- **Leonard Bolc (1934–2013)**
specjalista w zakresie sztucznej inteligencji

Lingwistyka w IPI PAN 2



Ale **skąd lingwistyka w IPI PAN?**



- **Leonard Bolc (1934–2013)**
specjalista w zakresie sztucznej inteligencji
- od 1987 w IPI PAN

Lingwistyka w IPI PAN 2



Ale **skąd lingwistyka w IPI PAN?**



- **Leonard Bolc** (1934–2013)
specjalista w zakresie sztucznej inteligencji
- od 1987 w IPI PAN, m.in. automatyczne
przetwarzanie języka naturalnego (NLP)

Lingwistyka w IPI PAN 2



Ale skąd lingwistyka w IPI PAN?



- **Leonard Bolc (1934–2013)**
specjalista w zakresie sztucznej inteligencji
- od 1987 w IPI PAN, m.in. automatyczne
przetwarzanie języka naturalnego (NLP)
- wtedy – **od poł. XX w. do pocz. XXI w.** – wydawało się, że
NLP **wymaga precyzyjnego i wysoce sformalizowanego**
opisu języka

Lingwistyka w IPI PAN 2



Ale skąd lingwistyka w IPI PAN?



- **Leonard Bolc (1934–2013)**
specjalista w zakresie sztucznej inteligencji
- od 1987 w IPI PAN, m.in. automatyczne
przetwarzanie języka naturalnego (NLP)
- wtedy – **od poł. XX w. do pocz. XXI w.** – wydawało się, że
NLP **wymaga precyzyjnego i wysoce sformalizowanego opisu języka**
- i takie **opisy aspektów polszczyzny** powstawały w zespole
prof. Bolca

Lingwistyka w IPI PAN 2



Ale skąd lingwistyka w IPI PAN?



- **Leonard Bolc (1934–2013)**
specjalista w zakresie sztucznej inteligencji
- od 1987 w IPI PAN, m.in. automatyczne **przetwarzanie języka naturalnego (NLP)**
- wtedy – **od poł. XX w. do pocz. XXI w.** – wydawało się, że NLP **wymaga precyzyjnego i wysoce sformalizowanego opisu języka**
- i takie **opisy aspektów polszczyzny** powstawały w zespole prof. Bolca, w tym prace doktorskie obronione w IPI PAN:
 - Mykowiecka 1999 (konstrukcje względne w formalizmie HPSG)
 - Kupść 2000 (klityki w formalizmie HPSG)
 - Marciniak 2001 (koreferencja zaimków w formalizmie HPSG)

Lingwistyka w IPI PAN 2



Ale skąd lingwistyka w IPI PAN?



- **Leonard Bolc** (1934–2013)
specjalista w zakresie sztucznej inteligencji
- od 1987 w IPI PAN, m.in. automatyczne **przetwarzanie języka naturalnego** (NLP)
- wtedy – **od poł. XX w. do pocz. XXI w.** – wydawało się, że NLP **wymaga precyzyjnego i wysoce sformalizowanego opisu języka**
- i takie **opisy aspektów polszczyzny** powstawały w zespole prof. Bolca, w tym prace doktorskie obronione w IPI PAN:
 - Mykowiecka 1999 (konstrukcje względne w formalizmie HPSG)
 - Kupść 2000 (klityki w formalizmie HPSG)
 - Marciniak 2001 (koreferencja zaimków w formalizmie HPSG)
 - Przepiórkowski 1999 (m.in. przypadek gramatyczny w HPSG; Uniwersytet w Tybindze, nostryfikacja w IPI PAN 2000)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga formalizmu.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), w większości na UW:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona w IPI PAN
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), w większości na UW:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona w IPI PAN
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW:**

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w formalizmach motywowanych lingwistycznie:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW:**

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w formalizmach motywowanych lingwistycznie:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW**:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w formalizmach motywowanych lingwistycznie:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga formalizmu.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW:**

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie:**

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW**:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW**:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW**:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW**:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 3



Sformalizowany opis polszczyzny wymaga **formalizmu**.

Pierwsze takie opisy – w **prologowym formalizmie gramatyk metamorficznych** (Colmerauer 1978, w tomie pod red. L. Bolca), **w większości na UW**:

- Szpakowicz 1978, 1986 – „mała gramatyka” na papierze z elementami implementacji
- Świdziński 1992 – „duża gramatyka” na papierze
- Bień 1997 – pierwsze wersje implementacji
- Woliński 2004 – pełna implementacja (i modyfikacja), praca obroniona **w IPI PAN**
- Ogrodniczuk 2006 – rozszerzenie i pewne wykorzystanie tej implementacji

Implementowalne opisy w IPI PAN – w **formalizmach motywowanych lingwistycznie**:

- **Head-driven Phrase Structure Grammar** (HPSG; Pollard i Sag 1987, 1994)
 - Przepiórkowski, Kupść, Marciniak, Mykowiecka 2002
„Formalny opis języka polskiego: Teoria i implementacja”
- **Lexical Functional Grammar** (LFG; Bresnan 1982)
 - Patejuk i Przepiórkowski 2017
„POLFIE: Współczesna gramatyka formalna języka polskiego” (oparta na obu powyższych)

Lingwistyka w IPI PAN 4



Przykład automatycznego rozbioru zdania na podstawie gramatyki POLFIE:

Fragment gramatyki:

```
IP -->
{
DEP IP
|
ILEX: ~=!
{
  @(TNS-EXISTS)
  @(MOOD-SPEC)
  |
  @(TNS-NOTEXISTS)
  {
    @(CAT-SUBC inf)
    {~(XCOMP ~)
      (~ SUBJ PRED)='pro'}
```

```
    |
    @(CAT-NEQ-PATH ~ inf)
  }
};
DEP*
|
IAUX VP
|
VP IAUX DEP*
}.
```

Ponad **dwa miliony wierszy kodu i komentarzy** (wliczając słownik; bez niego niecałe 20 tysięcy).

Lingwistyka w IPI PAN 4

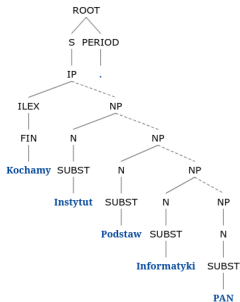


Przykład automatycznego rozbioru zdania na podstawie gramatyki POLFIE:

Fragment gramatyki:

```
IP -->
{
  DEP IP
  |
  ILEX: ~=!
  {
    @(TNS-EXISTS)
    @(MOOD-SPEC)
    |
    @(TNS-NOTEXISTS)
    {
      @(CAT-SUBC inf)
      {~(XCOMP ~)
      (~ SUBJ PRED)='pro'}
      |
      @(CAT-NEQ-PATH ~ inf)
    }
  }
};
DEP*
|
IAUX VP
|
VP IAUX DEP*
}.
```

C-structure

Ponad **dwa miliony wierszy kodu i komentarzy** (wliczając słownik; bez niego niecałe 20 tysięcy).

Lingwistyka w IPI PAN 4



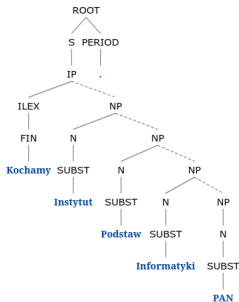
Przykład automatycznego rozbioru zdania na podstawie gramatyki POLFIE:

Fragment gramatyki:

```

IP -->
{
  DEP IP
  |
  ILEX: ~=!
  {
    @(TNS-EXISTS)
    @(MOOD-SPEC)
    |
    @(TNS-NOTEXISTS)
    {
      @(CAT-SUBC inf)
      {~(XCOMP ~)
      (~ SUBJ PRED)='pro'}
      |
      @(CAT-NEQ-PATH ~ inf)
    }
  }
};
DEP*
|
IAUX VP
|
VP IAUX DEP*
}.
```

C-structure



F-structure

PRED	'kochać<[126:pro], [116:Instytut]>'
TNS-ASP	127 TENSE pres, MOOD indicative, ASPECT imperf
PRED	'Instytut'
NTYPE	37 NSEM 38 COMMON count NSYN common
PRED	'podstawa'
NTYPE	207 NSEM 210 COMMON count NSYN common
PRED	'Informatyka'
NTYPE	110 NSEM 111 COMMON count NSYN common
PRED	'PAN'
POSS	NTYPE 7 NSEM 8 PROPER 9 PROPER-TYPE organization NSYN proper
PERS	47 PERS 3, NUM sg, GEND f, CASE gen
PERS	206 PERS 3, NUM sg, GEND f, CASE gen
PERS	41 PERS 3, NUM pl, GEND f, CASE gen
PERS	116 PERS 3, CASE acc, NUM sg, GEND m3
PRED	'pro'
SUBJ	126 PERS 1, NUM pl, CASE nom

Ponad **dwa miliony wierszy kodu i komentarzy** (wliczając słownik; bez niego niecałe 20 tysięcy).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki** są **potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza** w **Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna** w **Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- **LFG i HPSG** (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- **LFG i HPSG** (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- **LFG i HPSG** (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- **LFG i HPSG** (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza** w Instytucie Języka Polskiego PAN;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna** w Instytucie Podstaw Informatyki PAN.

Obecnie w IPI PAN:

- LFG i HPSG (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- LFG i HPSG (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- LFG i HPSG (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- **LFG i HPSG** (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka w IPI PAN 5



Przełom XX/XXI w.:

- jeszcze wydawało się, że takie **gramatyki są potrzebne do NLP**,
- ale już sukcesy zaczęły odnosić **systemy NLP oparte na statystyce i maszynowym uczeniu**.

Lata 10. XXI w.:

- dominują metody **maszynowego uczenia**;
- **wiedza lingwistyczna wykorzystywana do tworzenia zasobów treningowych NLP**.

Symptomatyczne:

- Patejuk 2015 (gramatyka konstrukcji współrzędnie złożonych w LFG) – obroniona jako **praca językoznawcza w Instytucie Języka Polskiego PAN**;
- Wróblewska 2014 (bank drzew zależnościowych i parser wytrenowany na nim) – obroniona jako **praca informatyczna w Instytucie Podstaw Informatyki PAN**.

Obecnie w IPI PAN:

- **LFG i HPSG** (i in. podejścia) są podstawą prac **teoretycznolingwistycznych**;
- z wykorzystaniem **tradycyjnego NLP** (parsery, słowniki elektroniczne).

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- zdominowane zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Lingwistyka i LLM-y 1



A **współczesne NLP**?

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Lingwistyka i LLM-y 1



A **współczesne NLP**?

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda?

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- zdominowane zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU)

Lingwistyka i LLM-y 1

A **współczesne NLP?**

- zdominowane zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- lata 80. XX w.: grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- Rumelhart i McClelland 1986 (monografia *Parallel Distributed Processing*)

Lingwistyka i LLM-y 1



A **współczesne NLP**?

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton

Lingwistyka i LLM-y 1



A **współczesne NLP**?

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987)

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987);
- **Elman 1990a,b**

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987);
- **Elman 1990a,b:**
 - psycholingwista z doktoratem lingwistycznym



Lingwistyka i LLM-y 1

A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987);
- **Elman 1990a,b:**
 - psycholingwista z doktoratem lingwistycznym,
 - współtwórca **rekurencyjnych sieci neuronowych** (Simple RNNs)

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987);
- **Elman 1990a,b:**
 - psycholingwista z doktoratem lingwistycznym,
 - współtwórca **rekurencyjnych sieci neuronowych** (Simple RNNs),
 - ważnych poprzedników architektury LLM-ów (Transformery)

Lingwistyka i LLM-y 1



A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987);
- **Elman 1990a,b:**
 - psycholingwista z doktoratem lingwistycznym,
 - współtwórca **rekurencyjnych sieci neuronowych** (Simple RNNs),
 - ważnych poprzedników architektury LLM-ów (Transformery),
 - jako pierwszy użył **zadania przewidywania kolejnego elementu sekwencji** (litery lub słowa) jako głównego celu treningowego dla prostej sieci rekurencyjnej



Lingwistyka i LLM-y 1

A **współczesne NLP?**

- **zdominowane** zostało przez **duże modele językowe** (LLM-y),
- które oczywiście nie mają **nic wspólnego z lingwistyką**.

Prawda? **Nie!**

Trochę historii:

- **lata 80. XX w.:** grupa badawcza *Parallel Distributed Processing* (UCSD / CMU);
- **Rumelhart i McClelland 1986** (monografia *Parallel Distributed Processing*):
 - rozdziały współautorstwa Geoffreya Hinton,
 - rozdziały o zastosowaniach językoznawczych (też Rumelhart i McClelland 1987);
- **Elman 1990a,b:**
 - psycholingwista z doktoratem lingwistycznym,
 - współtwórca **rekurencyjnych sieci neuronowych** (Simple RNNs),
 - ważnych poprzedników architektury LLM-ów (Transformery),
 - jako pierwszy użył **zadania przewidywania kolejnego elementu sekwencji** (litery lub słowa) jako głównego celu treningowego dla prostej sieci rekurencyjnej,
 - m.in. w zastosowaniach lingwistycznych (m.in. wyłanianie się klas gramatycznych).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale...

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? Zero, nothing.”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data...” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale...

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? Zero, nothing.”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data...” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale...

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? Zero, nothing.”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data...” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale...

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? **Zero, nothing.**”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data...” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale...

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? **Zero, nothing.**”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data...” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale. . .

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? **Zero, nothing.**”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data. . .” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

Lingwistyka i LLM-y 2



Lingwiści powinni być zachwyceni sukcesem LLM-ów? Wielu jest, ale...

Chomsky w wywiadach i korespondencji:

- „Does it tell you anything about human language? **Zero, nothing.**”
(2019, o uczeniu głębokim, ang. *deep learning*),
- „[It has some] ability to mimic some regularities in data...” (2022, o GPT-3),
- „In fact, its only achievement is to use up a lot of California’s energy” (jw.).

The New York Times

OPINION
GUEST ESSAY

Noam Chomsky: The False Promise of ChatGPT

March 8, 2023

„However **useful** these programs may be **in some narrow domains...**, we know from the science of linguistics and the philosophy of knowledge that **they differ profoundly from how humans reason and use language.**”

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego

– **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „ubóstwa bodźca” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „ubóstwa bodźca” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj 50–100 milionów słów (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. 300 miliardach tokenów (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. 15 bilionach tokenów (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Lingwistyka i LLM-y 3



Skąd taka reakcja Chomskiego? LLM-y osłabiły jedno z kluczowych twierdzeń Chomskiego – **natywizm**, według którego:

- człowiek rodzi się z genetycznie uwarunkowaną *specyficznie językową* zdolnością do posługiwania się językiem – tzw. **Gramatyką Uniwersalną**,
- co wyjaśnia, dlaczego dzieci opanowują złożone reguły językowe, także dotyczące zjawisk bardzo rzadkich – to argument z „**ubóstwa bodźca**” (ang. *poverty of stimulus*).

Problem: LLM-y uczą się języka bez wbudowanej Gramatyki Uniwersalnej.

To osłabia, ale **jeszcze nie unieważnia natywizmu**:

- **dzieci** słyszą średnio 2–7 milionów słów rocznie,
- do osiągnięcia pełnej płynności w wieku nastoletnim słyszą (lub czytają) w sumie zazwyczaj **50–100 milionów słów** (do 10^8 słów);
- **GPT-3** został wytrenowany na ok. **300 miliardach tokenów** (ponad 10^{11} słów – tysiąc razy więcej niż dziecko),
- **nowsze modele** (np. Llama 3) trenowane są na ok. **15 bilionach tokenów** (ponad 10^{13} słów – ponad sto tysięcy razy więcej niż dziecko; dojrzewanie musiałoby trwać ponad milion lat. . .).

Odpowiedź:



BabyLM Challenge

Sample-efficient pretraining on a developmentally plausible corpus

<https://babylm.github.io/>
(Warstadt i in. 2023)

Lingwistyka i LLM-y 4



Czy LLM-y są też problemem dla lingwistyki formalnej (uprawianej w IPI PAN)?

- „After decades of privilege and prominence in linguistics, **Noam Chomsky’s approach to the science of language is experiencing a remarkable downfall.**” **Ale:**
- „**[LLMs] appear to have representations of hierarchy...** and recursion, in the sense that they know about e.g. embedded sentences...”
- „They also almost certainly **have analogs of constraints, popular in approaches like [HPSG and LFG].**”

Lingwistyka i LLM-y 4



Czy LLM-y są też problemem dla lingwistyki formalnej (uprawianej w IPI PAN)?
Nie!

- „After decades of privilege and prominence in linguistics, **Noam Chomsky’s approach to the science of language is experiencing a remarkable downfall.**” **Ale:**
- „**[LLMs] appear to have representations of hierarchy...** and recursion, in the sense that they know about e.g. embedded sentences...”
- „They also almost certainly **have analogs of constraints, popular in approaches like [HPSG and LFG].**”

Lingwistyka i LLM-y 4



Czy LLM-y są też problemem dla lingwistyki formalnej (uprawianej w IPI PAN)?

Nie! Przyznają to nawet zagorzali przeciwnicy Chomskiego (Piantadosi 2024).

Modern language models refute Chomsky's approach to language

Steven T. Piantadosi

UC Berkeley & Helen Wills Neuroscience Institute

- „After decades of privilege and prominence in linguistics, **Noam Chomsky's approach to the science of language is experiencing a remarkable downfall.**” **Ale:**
- „**[LLMs] appear to have representations of hierarchy...** and recursion, in the sense that they know about e.g. embedded sentences...”
- „They also almost certainly **have analogs of constraints, popular in approaches like [HPSG and LFG].**”

Lingwistyka i LLM-y 4



Czy LLM-y są też problemem dla lingwistyki formalnej (uprawianej w IPI PAN)?

Nie! Przyznają to nawet zagorzali przeciwnicy Chomskiego (Piantadosi 2024).

Modern language models refute Chomsky's approach to language

Steven T. Piantadosi

UC Berkeley & Helen Wills Neuroscience Institute

- „After decades of privilege and prominence in linguistics, **Noam Chomsky's approach to the science of language is experiencing a remarkable downfall.**” **Ale:**
- „[LLMs] appear to have representations of hierarchy... and recursion, in the sense that they know about e.g. embedded sentences...”
- „They also almost certainly have analogs of constraints, popular in approaches like [HPSG and LFG].”

Lingwistyka i LLM-y 4



Czy LLM-y są też problemem dla lingwistyki formalnej (uprawianej w IPI PAN)?

Nie! Przyznają to nawet zagorzali przeciwnicy Chomskiego (Piantadosi 2024).

**Modern language models refute
Chomsky's approach to language**

Steven T. Piantadosi

UC Berkeley & Helen Wills Neuroscience Institute

- „After decades of privilege and prominence in linguistics, **Noam Chomsky's approach to the science of language is experiencing a remarkable downfall.**” **Ale:**
- „**[LLMs] appear to have representations of hierarchy...** and recursion, in the sense that they know about e.g. embedded sentences...”
- „They also almost certainly **have analogs of constraints, popular in approaches like [HPSG and LFG].**”

Lingwistyka i LLM-y 4



Czy LLM-y są też problemem dla lingwistyki formalnej (uprawianej w IPI PAN)?

Nie! Przyznają to nawet zagorzali przeciwnicy Chomskiego (Piantadosi 2024).

Modern language models refute Chomsky's approach to language

Steven T. Piantadosi

UC Berkeley & Helen Wills Neuroscience Institute

- „After decades of privilege and prominence in linguistics, **Noam Chomsky's approach to the science of language is experiencing a remarkable downfall.**” **Ale:**
- „**[LLMs] appear to have representations of hierarchy...** and recursion, in the sense that they know about e.g. embedded sentences...”
- „They also almost certainly **have analogs of constraints, popular in approaches like [HPSG and LFG].**”



Lingwistyka i LLM-y 5

Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and love the language models

Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodery Transformera)

Lingwistyka i LLM-y 5



Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and love the language models

Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodery Transformera)

Lingwistyka i LLM-y 5



Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and
love the language models

Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodér Transformera)



Lingwistyka i LLM-y 5

Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and love the language models

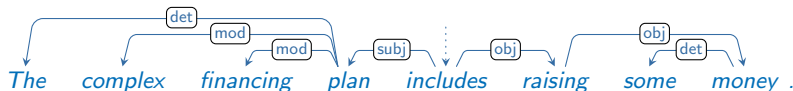
Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodery Transformera)

Dygresja – przykład struktury zależnościowej:



Lingwistyka i LLM-y 5



Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and love the language models

Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodery Transformera)

Dygresja – przykład struktury zależnościowej bez etykiet:



Lingwistyka i LLM-y 5



Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and love the language models

Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodler Transformera)

Dygresja – przykład struktury zależnościowej bez etykiet, graf nieskierowany:



Lingwistyka i LLM-y 5



Obecnie coraz bardziej prawdziwy jest tytuł pracy (Futrell i Mahowald 2026):

How linguistics learned to stop worrying and love the language models

Richard Futrell^a and Kyle Mahowald^b 

^aUniversity of California Irvine, Irvine, CA, USA, and ^bThe University of Texas at Austin, Austin, TX, USA

„[Many] studies have revealed increasing evidence of grammatical structure in models.”

Jedna z pierwszych takich prac – Hewitt i Manning 2019 – odkryła **struktury zależnościowe już we wcześniejszych architekturach** (ELMo – RNN, BERT – enkoder Transformera; obecne LLM-y – autoregresywny dekodler Transformera):

BERTlarge16

The complex financing plan in the S+L bailout law includes raising \$ 30 billion from debt issued by the newly created RTC .

Lingwistyka i LLM-y 6



Badania pokazują więc, że:

- reprezentacje zdań w LLM-ach kodują struktury gramatyczne,
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała...);
- czy ta gramatyka jest podobna do gramatyki języka naturalnego?

Lingwistyka i LLM-y 6



Badania pokazują więc, że:

- **reprezentacje zdań w LLM-ach kodują struktury gramatyczne,**
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała. . .);
- czy ta gramatyka jest podobna do gramatyki języka naturalnego?

Lingwistyka i LLM-y 6



Badania pokazują więc, że:

- **reprezentacje zdań w LLM-ach kodują struktury gramatyczne,**
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała...);
- czy ta gramatyka jest podobna do gramatyki języka naturalnego?

Lingwistyka i LLM-y 6



Badania pokazują więc, że:

- **reprezentacje zdań w LLM-ach kodują struktury gramatyczne,**
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała. . .);
- **czy ta gramatyka jest podobna do gramatyki języka naturalnego?**

Lingwistyka i LLM-y 6



Badania pokazują więc, że:

- **reprezentacje zdań w LLM-ach kodują struktury gramatyczne,**
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała. . .);
- **czy ta gramatyka jest podobna do gramatyki języka naturalnego?**

Warstadt i in. 2020:

BLiMP: The Benchmark of Linguistic Minimal Pairs for English

Alex Warstadt¹, Alicia Parrish¹, Haokun Liu², Anhad Mohananey²,
Wei Peng², Sheng-Fu Wang¹, Samuel R. Bowman^{1,2,3}

¹Department of Linguistics
New York University

²Department of Computer Science
New York University

³Center for Data Science
New York University

Lingwistyka i LLM-y 6



Badania pokazują więc, że:

- **reprezentacje zdań w LLM-ach kodują struktury gramatyczne,**
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała. . .);
- **czy ta gramatyka jest podobna do gramatyki języka naturalnego?**

Warstadt i in. 2020:

BLiMP: The Benchmark of Linguistic Minimal Pairs for English

Alex Warstadt¹, Alicia Parrish¹, Haokun Liu², Anhad Mohananey²,
Wei Peng², Sheng-Fu Wang¹, Samuel R. Bowman^{1,2,3}

¹Department of Linguistics
New York University

²Department of Computer Science
New York University

³Center for Data Science
New York University

Minimalne pary typu:

- Jeffrey **hasn't** criticized Donald.
- Jeffrey **haven't** criticized Donald.



Lingwistyka i LLM-y 6

Badania pokazują więc, że:

- reprezentacje zdań w LLM-ach kodują struktury gramatyczne,
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała...);
- czy ta gramatyka jest podobna do gramatyki języka naturalnego?

Warstadt i in. 2020:

BLiMP: The Benchmark of Linguistic Minimal Pairs for English

Alex Warstadt¹, Alicia Parrish¹, Haokun Liu², Anhad Mohananey²,
Wei Peng², Sheng-Fu Wang¹, Samuel R. Bowman^{1,2,3}

¹Department of Linguistics
New York University

²Department of Computer Science
New York University

³Center for Data Science
New York University

Minimalne pary typu:

- Jeffrey **hasn't** criticized Donald.
- Jeffrey **haven't** criticized Donald.
- Boys that aren't disturbing Natalie **suffer**.
- Boys that aren't disturbing Natalie **suffers**.



Lingwistyka i LLM-y 6

Badania pokazują więc, że:

- reprezentacje zdań w LLM-ach kodują struktury gramatyczne,
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała...);
- czy ta gramatyka jest podobna do gramatyki języka naturalnego?

Warstadt i in. 2020:

BLiMP: The Benchmark of Linguistic Minimal Pairs for English

Alex Warstadt¹, Alicia Parrish¹, Haokun Liu², Anhad Mohananey²,
Wei Peng², Sheng-Fu Wang¹, Samuel R. Bowman^{1,2,3}

¹Department of Linguistics
New York University

²Department of Computer Science
New York University

³Center for Data Science
New York University

Minimalne pary typu:

- Jeffrey **hasn't** criticized Donald.
- Jeffrey **haven't** criticized Donald.
- Boys that aren't disturbing Natalie **suffer**.
- Boys that aren't disturbing Natalie **suffers**.
- 12 zjawisk (tu: uzgodnienie podmiotu z czasownikiem)



Lingwistyka i LLM-y 6

Badania pokazują więc, że:

- **reprezentacje zdań w LLM-ach kodują struktury gramatyczne,**
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała...);
- **czy ta gramatyka jest podobna do gramatyki języka naturalnego?**

Warstadt i in. 2020:

BLiMP: The Benchmark of Linguistic Minimal Pairs for English

Alex Warstadt¹, Alicia Parrish¹, Haokun Liu², Anhad Mohananey²,
Wei Peng², Sheng-Fu Wang¹, Samuel R. Bowman^{1,2,3}

¹Department of Linguistics
New York University

²Department of Computer Science
New York University

³Center for Data Science
New York University

Minimalne pary typu:

- Jeffrey **hasn't** criticized Donald.
- Jeffrey **haven't** criticized Donald.
- Boys that aren't disturbing Natalie **suffer**.
- Boys that aren't disturbing Natalie **suffers**.
- 12 zjawisk (tu: uzgodnienie podmiotu z czasownikiem),
- kilka paradygmatów dla każdego (tu dwa), w sumie 67



Lingwistyka i LLM-y 6

Badania pokazują więc, że:

- reprezentacje zdań w LLM-ach kodują struktury gramatyczne,
- a więc **język LLM-ów ma gramatykę** (ale Eliza z lat 60. XX w. też jakąś miała...);
- czy ta gramatyka jest podobna do gramatyki języka naturalnego?

Warstadt i in. 2020:

BLiMP: The Benchmark of Linguistic Minimal Pairs for English

Alex Warstadt¹, Alicia Parrish¹, Haokun Liu², Anhad Mohananey²,
Wei Peng², Sheng-Fu Wang¹, Samuel R. Bowman^{1,2,3}

¹Department of Linguistics
New York University

²Department of Computer Science
New York University

³Center for Data Science
New York University

Minimalne pary typu:

- Jeffrey **hasn't** criticized Donald.
- Jeffrey **haven't** criticized Donald.
- Boys that aren't disturbing Natalie **suffer**.
- Boys that aren't disturbing Natalie **suffers**.
- 12 zjawisk (tu: uzgodnienie podmiotu z czasownikiem),
- kilka paradygmatów dla każdego (tu dwa), w sumie 67,
- 1000 par w paradygmacie (sztucznie utworzone, wg wzorców, niekiedy nienaturalne semantycznie).

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo** gramatycznej wersji zdania?

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo** gramatycznej wersji zdania?

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo gramatycznej wersji zdania?**

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Do utworzenia takich zbiorów testowych **potrzebna jest wiedza lingwistyczna...**

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo gramatycznej wersji zdania?**

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Do utworzenia takich zbiorów testowych **potrzebna jest wiedza lingwistyczna...**

Podobne zbiory testowe powstały dla parunastu innych języków, w tym:

- **rosyjskiego** (RuBLiMP; Taktasheva i in. 2024)

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo gramatycznej wersji zdania?**

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Do utworzenia takich zbiorów testowych **potrzebna jest wiedza lingwistyczna...**

Podobne zbiory testowe powstały dla parunastu innych języków, w tym:

- **rosyjskiego** (RuBLiMP; Taktasheva i in. 2024):
 - minimalne pary utworzone na podstawie korpusu drzew zależnościowych

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo gramatycznej wersji zdania?**

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Do utworzenia takich zbiorów testowych **potrzebna jest wiedza lingwistyczna...**

Podobne zbiory testowe powstały dla parunastu innych języków, w tym:

- **rosyjskiego** (RuBLiMP; Taktasheva i in. 2024):
 - minimalne pary utworzone na podstawie korpusu drzew zależnościowych;
- i **tureckiego** (TurBLiMP; Başar i in. 2025)

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo gramatycznej wersji zdania?**

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Do utworzenia takich zbiorów testowych **potrzebna jest wiedza lingwistyczna...**

Podobne zbiory testowe powstały dla parunastu innych języków, w tym:

- **rosyjskiego** (RuBLiMP; Taktasheva i in. 2024):
 - minimalne pary utworzone na podstawie korpusu drzew zależnościowych;
- i **tureckiego** (TurBLiMP; Başar i in. 2025):
 - minimalne pary utworzone (pół)ręcznie

Lingwistyka i LLM-y 7



Czy model przypisuje **większe prawdopodobieństwo gramatycznej wersji zdania?**

Model	Overall	ANA. AGR	ARG. STR	BINDING	CTRL. RAIS.	D-N AGR	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	NPI	QUANTIFIERS	S-V AGR
5-gram	61.2	47.9	71.9	64.4	68.5	70.0	36.9	60.2	79.5	57.2	45.5	53.5	60.3
LSTM	69.8	91.7	73.2	73.5	67.0	85.4	67.6	73.9	89.1	46.6	51.7	64.5	80.1
TXL	69.6	94.1	72.2	74.7	71.5	83.0	77.2	66.6	78.2	48.4	55.2	69.3	76.0
GPT-2	83.0	99.3	81.8	80.9	81.9	95.8	89.3	81.3	91.9	72.7	76.8	79.0	86.4
Human	88.6	97.5	90.0	87.3	83.9	92.2	85.0	86.9	97.0	84.9	88.1	86.6	90.9

Do utworzenia takich zbiorów testowych **potrzebna jest wiedza lingwistyczna...**

Podobne zbiory testowe powstały dla parunastu innych języków, w tym:

- **rosyjskiego** (RuBLiMP; Taktasheva i in. 2024):

- minimalne pary utworzone na podstawie korpusu drzew zależnościowych;

- i **tureckiego** (TurBLiMP; Başar i in. 2025):

- minimalne pary utworzone (pół)ręcznie;

- ale **nie dla polskiego**

(oprócz polskiego komponentu wielojęzycznego MultiBLiMP; Jumelet i in. 2026).



Lingwistyka i LLM-y 8

PolBLiMP?

- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I okazało się, że doniósł.
 - I okazał się, że doniósł.

Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PolBLiMP: Evaluating LLMs' mastery of subject–verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.



Lingwistyka i LLM-y 8

PolBLiMP?

- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I okazało się, że doniósł.
 - I okazał się, że doniósł.

Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PolBLiMP: Evaluating LLMs' mastery of subject-verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.



Lingwistyka i LLM-y 8

PoIBLiMP?

- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I okazało się, że doniósł.
 - I okazał się, że doniósł.

LINDAT Catalog

> > > Universal Dependencies 2.18 - Polish - LFG My q

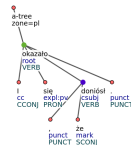
Relations Node Types Attributes Operators F

```
a-node $root := [  
  tag = 'VERB',  
  member iset [  
    number = 'sing',  
    gender = 'neut',  
  ],  
  child a-node $csubj := [  
    tag = 'VERB',  
    deprel = 'subj',  
    # exclude sentences with lemma 'kto' - incorrect annotations  
    &x child [lemma = 'kto']  
  ]  
]
```

Execute query T w/o Filters Suggest (0)

Result: 1 / 48

[pl] I okazało się, że doniósł.



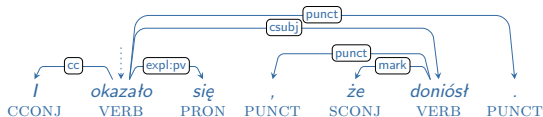
Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PoIBLiMP: Evaluating LLMs' mastery of subject-verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.

Lingwistyka i LLM-y 8



PoIBLiMP?

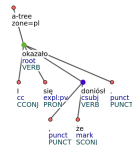
- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I okazało się, że doniósł.
 - I okazał się, że doniósł.



Screenshot of the LINDAT Universal Dependencies 2.18 - Polish - LFG interface. The query is:

```
a-node $root := [
  tag = 'VERB',
  member iset [
    number = 'sing',
    gender = 'neut'
  ],
]
```

The results show the sentence: [pl] I okazało się, że doniósł.



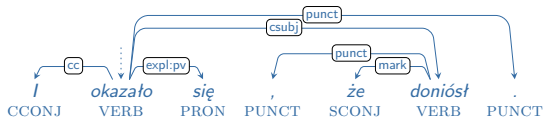
Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PoIBLiMP: Evaluating LLMs' mastery of subject-verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.

Lingwistyka i LLM-y 8



PoBLiMP?

- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I okazało się, że doniósł.
 - I okazał się, że doniósł.



LINDAT Catalog

Universal Dependencies 2.18 - Polish - LFG

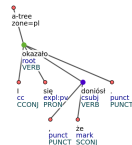
Relations Node Types Attributes Operators F

```
a-node $root := [  
  tag = 'VERB',  
  member iset [  
    number = 'sing',  
    gender = 'neut'  
  ],  
]  
  
child a-node $csubj := [  
  tag = 'VERB',  
  deprel = 'subj',  
  # exclude sentences with lemma 'kto' - incorrect annotations  
  &x child [lemma = 'kto']  
]  
]
```

Execute query w/o Filters Suggest (0)

Result: 1 / 48

[pl] I okazało się, że doniósł.



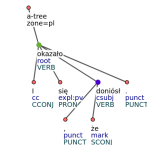
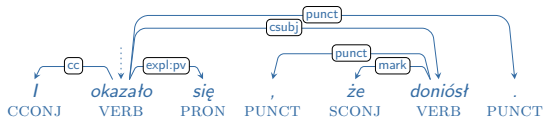
Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PoBLiMP: Evaluating LLMs' mastery of subject-verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.

Lingwistyka i LLM-y 8



PoIBLiMP?

- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I **okazało** się, że doniósł.
 - I **okazał** się, że doniósł.



LINDAT Catalog

Universal Dependencies 2.18 - Polish - LFG

Relations Node Types Attributes Operators F

```
a-node $root := [  
  tag = 'VERB',  
  member iset [  
    number = 'sing',  
    gender = 'neut',  
  ],  
  child a-node $csubj := [  
    tag = 'VERB',  
    deprel = 'subj',  
    # exclude sentences with lemma 'kto' - incorrect annotations  
    &x child [lemma = 'kto']  
  ]  
]
```

Execute query T w/o Filters Suggest (0)

Result: 1 / 48

[pl] I okazało się, że doniósł.

Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PoIBLiMP: Evaluating LLMs' mastery of subject-verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.

Lingwistyka i LLM-y 8



PoIBLiMP?

- automatyczne wyszukiwanie przykładów w polskich korpusach zależnościowych:
 - PDB-UD (Wróblewska 2014, 2018, 2020)
 - LFG Enhanced UD (Patejuk i Przepiórkowski 2018, Przepiórkowski i Patejuk 2020)
- na podstawie ręcznie skonstruowanych zapytań
- tworzenie niegramatycznego odpowiednika za pomocą słownika PoliMorf (Woliński i in. 2012)
- - I okazało się, że doniósł.
 - I okazał się, że doniósł.

● syntax_type	bielik_1.5	bielik_1.5_inst	bielik_7.0	bielik_7.0_inst	plum_8.0	plum_8.0_inst
subj_adjectival	0.9704	0.9807	0.9892	0.9805	0.9909	0.9879
subj_verb_gender	0.9426	0.9563	0.9883	0.9778	0.9910	0.9856
subj_verb_number	0.9527	0.9711	0.9919	0.9716	0.9861	0.9851
subj_verb_person	0.9816	0.9897	0.9924	0.9859	0.9935	0.9922

LINDAT Catalog

Universal Dependencies 2.18 - Polish - LFG My q

Relations Node Types Attributes Operators F

```

a-node $root := [
  tag = 'VERB',
  member iset [
    number = 'sing',
    gender = 'neut'
  ],
]

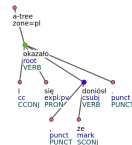
child a-node $subj := [
  tag = 'VERB',
  deprel = 'subj',
  # exclude sentences with lemma 'kto' - incorrect annotations
  &x child [lemma = 'kto']
]

```

Execute query T w/o Filters Suggest (0)

Result: 1 / 48

[pl] I okazało się, że doniósł.



Jakubowski A., Kędra H., Kwiatkowski M., Ustrzycki A., Przepiórkowski A., 2026: *Towards PoIBLiMP: Evaluating LLMs' mastery of subject-verb agreement in Polish*, zgłoszenie przyjęte na konferencję „Język a duże modele językowe: Oddziaływanie, badanie, przekład, dydaktyka”, 17–18 listopada 2026, Kielce.

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 9



Wnioski z tego typu ewaluacji (dla różnych języków):

- **LLM-y trafnie rozpoznają kontrasty gramatyczne,**
- w większości wypadków **na poziomie porównywalnym z ludzkim.**

To pokazuje, że **język LLM-ów jest zbliżony do języka ludzi**, ale **takie testy są:**

- **binarne** (czy zdanie gramatyczne jest oceniane lepiej niż niegramatyczne?),
- i **bardzo zgrubne** (czy oceniane jest choć odrobinę lepiej?).

U ludzi, w wypadku bardziej subtelnych zjawisk:

- **osądy gramatyczności** wynikające z testów psycholingwistycznych **nie są binarne**,
- i **nie są zgrubne** – dają stabilną „wielkość efektu”.

Czy tak samo jest z LLM-ami?

Lingwistyka i LLM-y 10

Przykład – **(nielokalny) dopełniacz negacji:**

● bez negacji – biernik:

- Czytam ten stary pamiętnik. (ACC)
- *Czytam tego starego pamiętnika. (GEN)

● z negacją – dopełniacz:

- *Nie czytam ten stary pamiętnik. (ACC)
- Nie czytam tego starego pamiętnika. (GEN)

● z negacją nielokalną nadal dopełniacz:

- *Nie mogę czytać ten stary pamiętnik. (ACC)
- Nie mogę czytać tego starego pamiętnika. (GEN)

● ale nie zawsze osądy są wyraziste:

- ? Nie żał czytać ten stary pamiętnik. (ACC)
- ? Nie żał czytać tego starego pamiętnika. (GEN)



Lingwistyka i LLM-y 10

Przykład – **(nielokalny) dopełniacz negacji:**

● **bez negacji – biernik:**

- Czytam ten stary pamiętnik. (ACC)
- *Czytam tego starego pamiętnika. (GEN)

● **z negacją – dopełniacz:**

- *Nie czytam ten stary pamiętnik. (ACC)
- Nie czytam tego starego pamiętnika. (GEN)

● **z negacją nielokalną nadal dopełniacz:**

- *Nie mogę czytać ten stary pamiętnik. (ACC)
- Nie mogę czytać tego starego pamiętnika. (GEN)

● **ale nie zawsze osądy są wyraziste:**

- ? Nie żał czytać ten stary pamiętnik. (ACC)
- ? Nie żał czytać tego starego pamiętnika. (GEN)

Lingwistyka i LLM-y 10

Przykład – **(nielokalny) dopełniacz negacji:**● **bez negacji – biernik:**

- Czytam ten stary pamiętnik. (ACC)
- *Czytam tego starego pamiętnika. (GEN)

● **z negacją – dopełniacz:**

- *Nie czytam ten stary pamiętnik. (ACC)
- Nie czytam tego starego pamiętnika. (GEN)

● **z negacją nielokalną nadal dopełniacz:**

- *Nie mogę czytać ten stary pamiętnik. (ACC)
- Nie mogę czytać tego starego pamiętnika. (GEN)

● **ale nie zawsze osądy są wyraziste:**

- ?Nie żał czytać ten stary pamiętnik. (ACC)
- ?Nie żał czytać tego starego pamiętnika. (GEN)



Lingwistyka i LLM-y 10

Przykład – **(nielokalny) dopełniacz negacji:**

- **bez negacji – biernik:**

- Czytam ten stary pamiętnik. (ACC)
- *Czytam tego starego pamiętnika. (GEN)

- **z negacją – dopełniacz:**

- *Nie czytam ten stary pamiętnik. (ACC)
- Nie czytam tego starego pamiętnika. (GEN)

- **z negacją nielokalną nadal dopełniacz:**

- *Nie mogę czytać ten stary pamiętnik. (ACC)
- Nie mogę czytać tego starego pamiętnika. (GEN)

- **ale nie zawsze osądy są wyraziste:**

- ? Nie żał czytać ten stary pamiętnik. (ACC)
- ? Nie żał czytać tego starego pamiętnika. (GEN)



Lingwistyka i LLM-y 10

Przykład – **(nielokalny) dopełniacz negacji:**

- **bez negacji – biernik:**

- Czytam ten stary pamiętnik. (ACC)
- *Czytam tego starego pamiętnika. (GEN)

- **z negacją – dopełniacz:**

- *Nie czytam ten stary pamiętnik. (ACC)
- Nie czytam tego starego pamiętnika. (GEN)

- **z negacją nielokalną nadal dopełniacz:**

- *Nie mogę czytać ten stary pamiętnik. (ACC)
- Nie mogę czytać tego starego pamiętnika. (GEN)

- **ale nie zawsze osądy są wyraziste:**

- ? Nie żał czytać ten stary pamiętnik. (ACC)
- ? Nie żał czytać tego starego pamiętnika. (GEN)



Lingwistyka i LLM-y 10

Przykład – **(nielokalny) dopełniacz negacji**:

- **bez negacji – biernik**:

- Czytam ten stary pamiętnik. (ACC)
- *Czytam tego starego pamiętnika. (GEN)

- **z negacją – dopełniacz**:

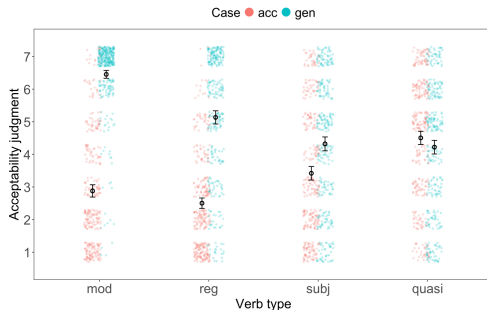
- *Nie czytam ten stary pamiętnik. (ACC)
- Nie czytam tego starego pamiętnika. (GEN)

- **z negacją nielokalną nadal dopełniacz (mod)**:

- *Nie mogę czytać ten stary pamiętnik. (ACC)
- Nie mogę czytać tego starego pamiętnika. (GEN)

- **ale nie zawsze osądy są wyraziste (quasi)**:

- ?Nie żał czytać ten stary pamiętnik. (ACC)
- ?Nie żał czytać tego starego pamiętnika. (GEN)



Lingwistyka i LLM-y 11

**Jak korelują osądy ludzi z osądami LLM-ów (mierzonymi zaskoczeniem; ang. *surprisal*)?**

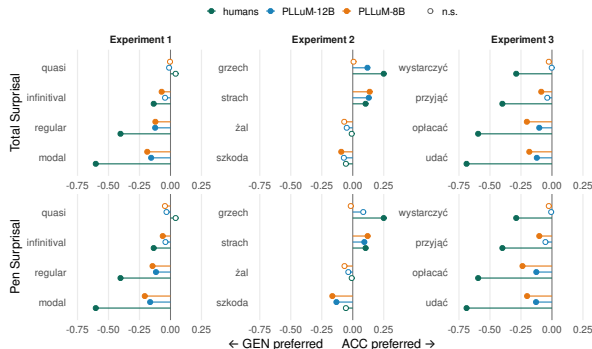
- jest widoczna korelacja,
- choć „wielkość efektu” jest mniejsza u LLM-ów,
- i siła korelacji nie jest wielka.



Lingwistyka i LLM-y 11

Jak korelują osądy ludzi z osądami LLM-ów (mierzonymi zaskoczeniem; ang. *surprisal*)?

- jest **widoczna korelacja**,
- choć „wielkość efektu” jest **mniejsza u LLM-ów**,
- i siła korelacji nie jest wielka.



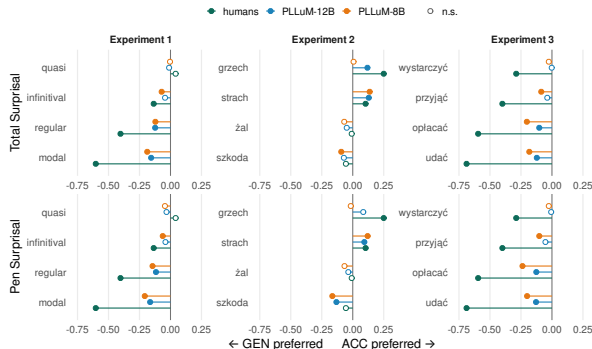
Şenşekerçi B., Gurak M., 2026: *How Surprising Is Unacceptable? Evidence from Polish Genitive of Negation*, prezentacja na konferencji „Architectures and Mechanisms for Language Processing” (AMLaP 2026), 2–4 września 2026, Saarbrücken.



Lingwistyka i LLM-y 11

Jak korelują osądy ludzi z osądami LLM-ów (mierzonymi zaskoczeniem; ang. *surprisal*)?

- jest **widoczna korelacja**,
- choć „**wielkość efektu**” jest **mniejsza u LLM-ów**,
- i **siła korelacji nie jest wielka**.



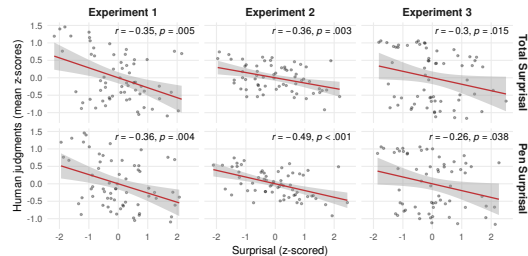
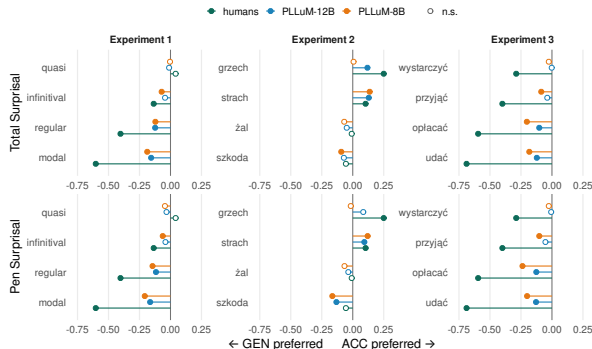
Şenşekerçi B., Gurak M., 2026: *How Surprising Is Unacceptable? Evidence from Polish Genitive of Negation*, prezentacja na konferencji „Architectures and Mechanisms for Language Processing” (AMLaP 2026), 2–4 września 2026, Saarbrücken.



Lingwistyka i LLM-y 11

Jak korelują osądy ludzi z osądami LLM-ów (mierzonymi zaskoczeniem; ang. *surprisal*)?

- jest widoczna korelacja,
- choć „wielkość efektu” jest mniejsza u LLM-ów,
- i siła korelacji nie jest wielka.



(Im mniej akceptowalne zdanie dla człowieka, tym większe „zaskoczenie” modelu – dlatego linie opadają).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale są **pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: **badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka** (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 12



Podsumowując ostatnie slajdy:

- **język LLM-ów jest bardzo zbliżony do języka ludzi** (BLiMP, eksperymenty),
- ale **są pewne różnice** (być może wynikające z wielojęzyczności LLM-ów).

Skoro tak, wyłania się **możliwe zastosowanie LLM-ów w językoznawstwie**:

- przeprowadzanie eksperymentów psycholingwistycznych na ludziach jest czasochłonne i kosztowne,
- więc przeprowadza się mniej eksperymentów, niż by się chciało;
- eksperymenty na LLM-ach są mniej obciążające,
- więc **można użyć LLM-ów do wstępnej oceny eksperymentów**
- i przeprowadzić na ludziach te, które najlepiej rokują (ale nie tylko te).

Inne zastosowanie: badanie różnic między gramatycznością i ograniczeniami w przetwarzaniu języka (tematy prac doktorskich B. Şenşekerci i M. Gurak; Międzydziedzinowa Szkoła Doktorska UW).

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, którego mój pies kiedyś pogryzł, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, którego twój kot gonił przez kwadrans, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, którego mój pies kiedyś pogryzł, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, którego twój kot gonił przez kwadrans, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, którego mój pies kiedyś pogryzł, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, którego twój kot gonił przez kwadrans, w końcu uciekł. (Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, którego mój pies kiedyś pogryzł, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, którego twój kot gonił przez kwadrans, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, którego mój pies kiedyś pogryzł, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: akceptowalność $\neq$ gramatyczność.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, którego mój pies kiedyś pogryzł, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł. (Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomym przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to zjawisko z **poziomu przetwarzania**, a nie gramatyczności.



Lingwistyka i LLM-y 13

Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to zjawisko z **poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to zjawisko z **poziomu przetwarzania**, a nie gramatyczności.



Lingwistyka i LLM-y 13

Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to zjawisko z **poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to zjawisko z **poziomu przetwarzania**, a nie gramatyczności.



Lingwistyka i LLM-y 13

Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to zjawisko z **poziomu przetwarzania**, a nie gramatyczności.

Lingwistyka i LLM-y 13



Problem: **akceptowalność** $\neq$ **gramatyczność**.

Przykład (ang. *center embedding*; Miller i Chomsky 1963):

- Szczur, którego twój kot, **którego mój pies kiedyś pogryzł**, gonił przez kwadrans, w końcu uciekł.
- Szczur w końcu uciekł. (Tego szczura gonił przez kwadrans twój kot).
- Szczur, **którego twój kot gonił przez kwadrans**, w końcu uciekł.
(Tego kota kiedyś pogryzł mój pies).

Powyższe (pierwsze) zdanie jest raczej **nieakceptowalne**, ale **dłaczego?**

- **niegramatyczne?**
- gramatyczne, ale **nieprzetwarzalne z powodu ograniczeń pamięci krótkotrwałej?**

Możliwe zastosowanie LLM-ów w językoznawstwie:

- osądy *gramatyczności* LLM-ów korelują z ludzkimi,
- ale **ich ograniczenia przetwarzania są (zapewne) zupełnie inne**,
- więc jeżeli ludzkie osądy *akceptowalności* jakiegoś zjawiska **nie korelują** z osądami LLM-ów, to być może jest to **zjawisko z poziomu przetwarzania**, a nie gramatyczności.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach.**
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

- minimalizacja zależności funkcyjnych (np. od przyimka) – gramatyka,
- minimalizacja zależności leksykalnych (wszystkich innych powyżej) – przetwarzanie.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

- minimalizacja zależności funkcyjnych (np. od przyimka) – gramatyka,
- minimalizacja zależności leksykalnych (wszystkich innych powyżej) – przetwarzanie.

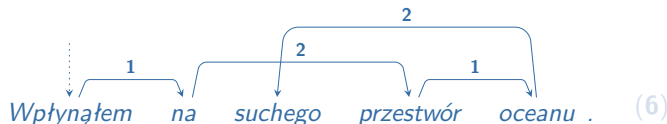
W przybliżeniu: **Preferuj** zdania o krótszych sumarycznie zależnościach.
(I lepiej, żeby te zależności się nie przecinały).



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

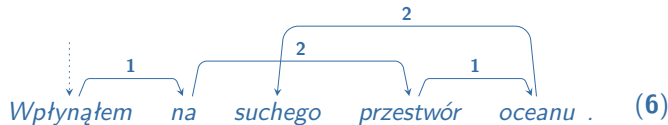
- minimalizacja zależności funkcyjnych (np. od przyimka) – gramatyka,
- minimalizacja zależności leksykalnych (wszystkich innych powyżej) – przetwarzanie.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

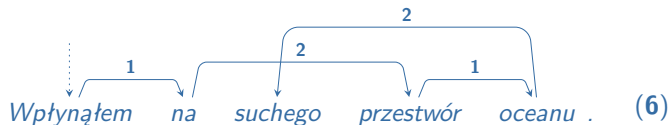
- minimalizacja zależności funkcyjnych (np. od przyimka) – gramatyka,
- minimalizacja zależności leksykalnych (wszystkich innych powyżej) – przetwarzanie.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

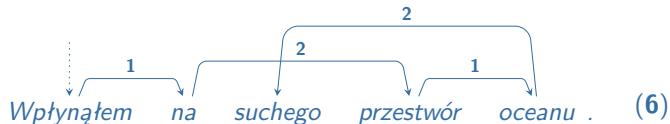
- minimalizacja zależności funkcyjnych (np. od przyimka) – gramatyka,
- minimalizacja zależności leksykalnych (wszystkich innych powyżej) – przetwarzanie.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

- minimalizacja zależności funkcyjnych (np. od przyimka) – gramatyka,
- minimalizacja zależności leksykalnych (wszystkich innych powyżej) – przetwarzanie.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

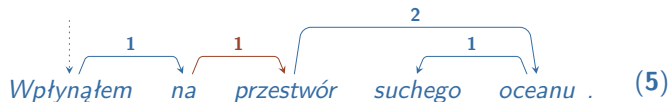
- **minimalizacja zależności funkcyjnych** (np. od przyimka) – **gramatyka**,
- **minimalizacja zależności leksykalnych** (wszystkich innych powyżej) – **przetwarzanie**.



Lingwistyka i LLM-y 14

Zjawisko, co do którego nie jest jasne, w jakim stopniu jest gramatyczne, to **minimalizacja długości zależności** (ang. *Dependency Length Minimization*; Gibson 1998).

W przybliżeniu: **Preferuj zdania o krótszych sumarycznie zależnościach**.
(I lepiej, żeby te zależności się nie przecinały).



Gerdes 2026:

- **minimalizacja zależności funkcyjnych** (np. od przyimka) – **gramatyka**,
- **minimalizacja zależności leksykalnych** (wszystkich innych powyżej) – **przetwarzanie**.

Lingwistyka i LLM-y 15



The Grammar Does the Work: Functional vs. Lexical Dependency Length Minimization Across Universal Dependencies

Kim Gerdes*

Université Paris-Saclay, LISN (CNRS)
Orsay, France

Lingwistyka i LLM-y 15



The Grammar Does the Work: Functional vs. Lexical Dependency Length Minimization Across Universal Dependencies

Kim Gerdes*
Université Paris-Saclay, LISN (CNRS)
Orsay, France

- **Argumentacja przekonująca,**
ale nie do końca konkluzywna

Lingwistyka i LLM-y 15



The Grammar Does the Work: Functional vs. Lexical Dependency Length Minimization Across Universal Dependencies

Kim Gerdes*

Université Paris-Saclay, LISN (CNRS)
Orsay, France

- **Argumentacja przekonująca,**
ale nie do końca konkluzywna,
- być może pomogłyby
eksperymenty na LLM-ach

The Grammar Does the Work: Functional vs. Lexical Dependency Length Minimization Across Universal Dependencies

Kim Gerdes*

Université Paris-Saclay, LISN (CNRS)
Orsay, France

- Argumentacja przekonująca,
ale nie do końca konkluzywna,
- być może pomogłyby
eksperymenty na LLM-ach,
- szczególnie skoro to **one napisały ten tekst!**

Ethics Statement and AI Disclosure

We fully disclose that this paper was produced using an AI system. The research idea, all analysis code, and the entire text were generated by Claude Opus 4.6 (Anthropic), operating in agentic mode via GitHub Copilot Chat in VS Code. The human author provided prompts and accepted or rejected proposed work at each step, but did not originate the central research question, write any code, or draft any prose beyond the prompts themselves. The author's contribution was verification and editorial oversight rather than generation of the core ideas, code, or prose.

The project was inspired by an experiment by D. Yanagizawa-Drott, who prompted an LLM to write a macroeconomics job market paper and publicly reflected on the implications (<https://x.com/YanagizawaD/status/2022034189395407093>). This paper began with an analogous prompt: *"I always dreamt of becoming a syntactician and a typologist one day. I want to submit a paper to the UDW workshop. Can you help me?"* The functional-lexical distinction that became the core contribution emerged during the AI's iterative exploration of the data, not from the human author. Prior work was identified using Google Scholar Labs and provided to the system for integration.

The reviews by anonymous UDW 2026 reviewers and by Sylvain Kahane were essential for improving the final version; reviewer feedback was given to the AI, which implemented the revisions. The code, analyses, and results were checked by the author, and the final paper benefited substantially from this review process.



Lingwistyka i LLM-y 15

The Grammar Does the Work: Functional vs. Lexical Dependency Length Minimization Across Universal Dependencies

Kim Gerdes*

Université Paris-Saclay, LISN (CNRS)
Orsay, France

- Argumentacja przekonująca, ale nie do końca konkluzywna,
- być może pomogłyby eksperymenty na LLM-ach,
- szczególnie skoro to **one napisały ten tekst!**

Kolejne zastosowanie LLM-ów w językoznawstwie: **wspomaganie pracy badawczej.**

Ethics Statement and AI Disclosure

We fully disclose that this paper was produced using an AI system. The research idea, all analysis code, and the entire text were generated by Claude Opus 4.6 (Anthropic), operating in agentic mode via GitHub Copilot Chat in VS Code. The human author provided prompts and accepted or rejected proposed work at each step, but did not originate the central research question, write any code, or draft any prose beyond the prompts themselves. The author's contribution was verification and editorial oversight rather than generation of the core ideas, code, or prose.

The project was inspired by an experiment by D. Yanagizawa-Drott, who prompted an LLM to write a macroeconomics job market paper and publicly reflected on the implications (<https://x.com/YanagizawaD/status/2022034189395407093>). This paper began with an analogous prompt: *"I always dreamt of becoming a syntactician and a typologist one day. I want to submit a paper to the UDW workshop. Can you help me?"* The functional-lexical distinction that became the core contribution emerged during the AI's iterative exploration of the data, not from the human author. Prior work was identified using Google Scholar Labs and provided to the system for integration.

The reviews by anonymous UDW 2026 reviewers and by Sylvain Kahane were essential for improving the final version; reviewer feedback was given to the AI, which implemented the revisions. The code, analyses, and results were checked by the author, and the final paper benefited substantially from this review process.



Lingwistyka i LLM-y 15

The Grammar Does the Work: Functional vs. Lexical Dependency Length Minimization Across Universal Dependencies

Kim Gerdes*

Université Paris-Saclay, LISN (CNRS)
Orsay, France

- Argumentacja przekonująca, ale nie do końca konkluzywna,
- być może pomogłyby eksperymenty na LLM-ach,
- szczególnie skoro to **one napisały ten tekst!**

Kolejne zastosowanie LLM-ów w językoznawstwie: **wspomaganie pracy badawczej.**

(Niniejsza prezentacja powstała przy
bardzo umiarkowanej pomocy LLM-ów. . .)

Ethics Statement and AI Disclosure

We fully disclose that this paper was produced using an AI system. The research idea, all analysis code, and the entire text were generated by Claude Opus 4.6 (Anthropic), operating in agentic mode via GitHub Copilot Chat in VS Code. The human author provided prompts and accepted or rejected proposed work at each step, but did not originate the central research question, write any code, or draft any prose beyond the prompts themselves. The author's contribution was verification and editorial oversight rather than generation of the core ideas, code, or prose.

The project was inspired by an experiment by D. Yanagizawa-Drott, who prompted an LLM to write a macroeconomics job market paper and publicly reflected on the implications (<https://x.com/YanagizawaD/status/2022034189395407093>). This paper began with an analogous prompt: *"I always dreamt of becoming a syntactician and a typologist one day. I want to submit a paper to the UDW workshop. Can you help me?"* The functional-lexical distinction that became the core contribution emerged during the AI's iterative exploration of the data, not from the human author. Prior work was identified using Google Scholar Labs and provided to the system for integration.

The reviews by anonymous UDW 2026 reviewers and by Sylvain Kahane were essential for improving the final version; reviewer feedback was given to the AI, which implemented the revisions. The code, analyses, and results were checked by the author, and the final paper benefited substantially from this review process.



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans... use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- **kodują struktury gramatyczne** w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans... use language”).

Językoznawstwo dla LLM-ów:

- **badanie języka LLM-ów:**
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans... use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans... use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).

Podsumowanie



Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).

Podsumowanie



Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).

Podsumowanie



Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- **badanie języka LLM-ów.**
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- **badanie języka LLM-ów:**
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).

Podsumowanie



Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).

Podsumowanie



Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans. . . use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).



Podsumowanie

Duże modele językowe

- kodują struktury gramatyczne w swoich reprezentacjach,
- odróżniają konstrukcje gramatyczne od niegramatycznych,
- ich „osądy gramatyczności” korelują z osądami ludzi
- (por. Chomsky i in. 2023: „[LLMs] differ profoundly from how humans... use language”).

Językoznawstwo dla LLM-ów:

- badanie języka LLM-ów:
 - konstrukcja zbiorów testowych typu BLiMP,
 - badanie struktur gramatycznych w reprezentacjach LLM-ów,
 - badania „przyczynowe” (ang. *causal*) – temat na inny referat.

LLM-y dla językoznawstwa:

- generowanie i wstępne testowanie hipotez językoznawczych,
- badanie relacji między gramatycznością i ograniczeniami w przetwarzaniu języka,
- wspomaganie pracy językoznawcy (na różnych etapach i w różnym zakresie).

Dziękuję za uwagę!

- Başar E., Padovani F., Jumelet J., Bisazza A., 2025: *TurBLiMP: A Turkish benchmark of linguistic minimal pairs*, [w:] C. Christodoulopoulos, T. Chakraborty, C. Rose, V. Peng (red.) *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Suzhou, China, s. 16495–16510, DOI: <https://doi.org/10.18653/v1/2025.emnlp-main.834>.
- Bień J.S., 1997: *Komputerowa weryfikacja formalnej gramatyki Świdzińskiego*, „Biuletyn Polskiego Towarzystwa Językoznawczego”, LII, s. 147–164.
- Bresnan J. (red.) 1982: *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, MA.
- Chomsky N., 1957: *Syntactic Structures*, Mouton, Haga.
- Chomsky N., 1965: *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.
- Chomsky N., 1981: *Lectures on Government and Binding*, Foris, Dordrecht.
- Chomsky N., 1986: *Barriers*, MIT Press, Cambridge, MA.
- Chomsky N., 1995: *The Minimalist Program*, MIT Press, Cambridge, MA.
- Chomsky N., 2001: *Derivation by phase*, [w:] M. Kenstowicz (red.) *Ken Hale: A Life in Language*, MIT Press, Cambridge, MA, s. 1–52, DOI: <https://doi.org/10.7551/mitpress/4056.003.0004>.
- Chomsky N., Roberts I., Watumull J., 2023: *Noam Chomsky: The false promise of ChatGPT*, The New York Times: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>.
- Colmerauer A., 1978: *Metamorphosis grammars*, [w:] L. Bolc (red.) *Natural Language Communications with Computers*, tom 63 z serii *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, s. 133–189.
- Elman J.L., 1990a: *Finding structure in time*, „Cognitive Science”, 14(2), s. 179–211.
- Elman J.L., 1990b: *Representation and structure in connectionist models*, [w:] G.T.M. Altmann (red.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, MIT Press, Cambridge, MA, s. 345–382.
- Futrell R., Mahowald K., 2026: *How linguistics learned to stop worrying and love the language models*, „Behavioral and Brain Sciences”, 49(e198), s. 1–80, DOI: <https://doi.org/10.1017/S0140525X2510112X>.

- Gerdes K., 2026: *The grammar does the work: Functional vs. lexical Dependency Length Minimization across Universal Dependencies*, [w:] *Proceedings of the Ninth Workshop on Universal Dependencies (UDW 2026)*, ELRA, s. 163–173.
- Gibson E., 1998: *Linguistic complexity: Locality of syntactic dependencies*, „Cognition”, 68(1), s. 1–76, DOI: 10.1016/s0010-0277(98)00034-1.
- Hewitt J., Manning C.D., 2019: *A structural probe for finding syntax in word representations*, [w:] J. Burstein, C. Doran, T. Solorio (red.) *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Association for Computational Linguistics, Minneapolis, Minnesota, s. 4129–4138, DOI: <https://doi.org/10.18653/v1/N19-1419>.
- Jumelet J., Weissweiler L., Nivre J., Bisazza A., 2026: *MultiBLiMP 1.0: A massively multilingual benchmark of linguistic minimal pairs*, „Transactions of the Association for Computational Linguistics”, 14, s. 193–216, DOI: 10.1162/tac1.a.600.
- Kupść A., 2000: *An HPSG Grammar of Polish Clitics*, Rozprawa doktorska, Institute of Computer Science, Polish Academy of Sciences, and Université Paris 7, Warszawa.
- Manning C.D., Clark K., Hewitt J., Khandelwal U., Levy O., 2020: *Emergent linguistic structure in artificial neural networks trained by self-supervision*, „Proceedings of the National Academy of Sciences”, 117(48), s. 30046–30054.
- Marciniak M., 2001: *Algorytmy implementacyjne syntaktycznych reguł koreferencji zaimków dla języka polskiego w terminach HPSG*, Rozprawa doktorska, Instytut Podstaw Informatyki PAN, Warszawa.
- Miller G.A., Chomsky N., 1963: *Finitary models of language users*, [w:] R.D. Luce, R.R. Bush, E. Galanter (red.) *Handbook of Mathematical Psychology*, tom II, John Wiley and Sons, Inc., New York and London, s. 419–491.
- Mykowiecka A., 1999: *Opis składniowy polskich konstrukcji względnych w formalizmie HPSG*, Rozprawa

- doktorska, Instytut Podstaw Informatyki PAN, Warszawa.
- Ogrodniczuk M., 2006: *Weryfikacja korpusu wypowiedników polskich (z wykorzystaniem gramatyki formalnej Świdzińskiego)*, Rozprawa doktorska, Wydział Neofilologii Uniwersytetu Warszawskiego, Warszawa.
- Patejuk A., 2015: *Unlike Coordination in Polish: An LFG Account*, Rozprawa doktorska, Instytut Języka Polskiego PAN, Kraków.
- Patejuk A., Przepiórkowski A., 2017: *POLFIE: współczesna gramatyka formalna języka polskiego*, „Język Polski”, XCVII(1), s. 48–64.
- Patejuk A., Przepiórkowski A., 2018: *From Lexical Functional Grammar to Enhanced Universal Dependencies: Linguistically Informed Treebanks of Polish*, Instytut Podstaw Informatyki PAN, Warszawa.
- Piantadosi S.T., 2024: *Modern language models refute Chomsky's approach to language*, [w:] *From Fieldwork to Linguistic Theory: A Tribute to Dan Everett*, Language Science Press, s. 353–414, DOI: <https://doi.org/10.5281/zenodo.12665933>.
- Pollard C., Sag I.A., 1987: *Information-Based Syntax and Semantics, Volume 1: Fundamentals*, CSLI Publications, Stanford, CA.
- Pollard C., Sag I.A., 1994: *Head-driven Phrase Structure Grammar*, University of Chicago 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., Mykowiecka A., 2002: *Formalny opis języka polskiego: Teoria i implementacja*, Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Przepiórkowski A., Patejuk A., 2020: *From Lexical Functional Grammar to enhanced Universal Dependencies: The UD-LFG treebank of Polish*, „Language Resources and Evaluation”, 54, s. 185–221, DOI: <https://doi.org/10.1007/s10579-018-9433-z>.
- Rumelhart D.E., McClelland J.L., 1986: *Parallel Distributed Processing*, MIT Press, Cambridge, MA.

- Rumelhart D.E., McClelland J.L., 1987: *Learning the past tenses of English verbs: Implicit rules or parallel distributed processing?*, [w:] B. MacWhinney (red.) *Mechanisms of Language Acquisition*, Psychology Press, New York, s. 195–248.
- 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, drugie wyd.
- Świdziński M., 1992: *Gramatyka formalna języka polskiego*, tom 349 z serii *Rozprawy Uniwersytetu Warszawskiego*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Taktasheva E., Bazhukov M., Koncha K., Fenogenova A., Artemova E., Mikhailov V., 2024: *RuBLiMP: Russian benchmark of linguistic minimal pairs*, [w:] Y. Al-Onaizan, M. Bansal, Y.N. Chen (red.) *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Miami, Florida, USA, s. 9268–9299, DOI: <https://doi.org/10.18653/v1/2024.emnlp-main.522>.
- Warstadt A., Mueller A., Choshen L., Wilcox E., Zhuang C., Ciro J., Mosquera R., Paranjape B., Williams A., Linzen T., Cotterell R., 2023: *Findings of the BabyLM challenge: Sample-efficient pretraining on developmentally plausible corpora*, [w:] *Proceedings of the BabyLM Challenge at the 27th Conference on Computational Natural Language Learning*, Association for Computational Linguistics, Singapore, s. 1–34, DOI: <https://doi.org/10.18653/v1/2023.conll-babylm.1>.
- Warstadt A., Parrish A., Liu H., Mohananey A., Peng W., Wang S.F., Bowman S.R., 2020: *BLiMP: The benchmark of linguistic minimal pairs for English*, „Transactions of the Association for Computational Linguistics”, 8, s. 377–392, DOI: https://doi.org/10.1162/tac1_a_00321.
- Woliński M., 2004: *Komputerowa weryfikacja gramatyki Świdzińskiego*, Rozprawa doktorska, Instytut Podstaw Informatyki PAN, Warszawa.

- Woliński M., Miłkowski M., Ogrodniczuk M., Przepiórkowski A., Szałkiewicz Ł., 2012: *PoliMorf: A (not so) new open morphological dictionary for Polish*, [w:] *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012*, European Language Resources Association (ELRA), Istanbul, Turkey, s. 860–864.
- Wróblewska A., 2014: *Polish Dependency Parser Trained on an Automatically Induced Dependency Bank*, Rozprawa doktorska, Instytut Podstaw Informatyki PAN, Warszawa.
- Wróblewska A., 2018: *Extended and enhanced Polish dependency bank in Universal Dependencies format*, [w:] M.C.d. Marneffe, T. Lynn, S. Schuster (red.) *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)*, Association for Computational Linguistics, s. 173–182.
- Wróblewska A., 2020: *Towards the Conversion of National Corpus of Polish to Universal Dependencies*, [w:] *Proceedings of the 12th Language Resources and Evaluation Conference*, European Language Resources Association (ELRA), Marseille, France, s. 5308–5315.

RuBLiMP 1



Taktasheva i in. 2024:

	Language	Size	# Paradigm	Method
BLiMP	English	67k	67	Dictionary & templates
CLiMP	Chinese	16k	16	Translation & templates
JBLiMP	Japanese	331	39	Extract from articles
SLING	Chinese	38k	38	UD Treebank & templates
NoCoLA_{zero}	Norwegian	99.1k	11	Extract from an L2 corpus
DaLAJ	Swedish	4.8k	4	Extract from an L2 corpus
LINDSEA	Indonesian	380	38	Expert-written min. pairs
	Tamil	200	20	
CLAMS	English	153.5k	13	Translation & templates
	French	49.3k	7	
	German	47.8k	7	
	Hebrew	40.8k	7	
	Russian	40.1k	7	
RuBLiMP	Russian	45k	45	Open text corpora, rules, automatic UD annotation, pretraining data detection

Table 1: Comparison of benchmarks of linguistic minimal pairs for different languages: BLiMP (Warstadt et al., 2020), CLiMP (Xiang et al., 2021), JBLiMP (Someya and Oseki, 2023), SLING (Song et al., 2022), NoCoLA_{zero} (Jentoft and Samuel, 2023), DaLAJ (Voldina et al., 2021), LINDSEA (Leong et al., 2023), CLAMS (Mueller et al., 2020), and RuBLiMP (ours).

RuBLiMP 2



Taktasheva i in. 2024:

Model	WORD FORM.	WORD INFL.	GOVERNMENT	SUBL.-PRED. AGR.	ANAPHOR AGR.	NP AGR.	FLOAT. QUANT. AGR.	REFLEXIVES	NEGATION	ARG. STRUCTURE	ASPECT	TENSE	AVERAGE
ruBERT-base	81.90	84.87	90.72	91.16	85.90	83.57	91.40	78.70	77.77	88.52	96.07	87.17	86.48
ruBERT-large	82.83	86.03	90.66	91.43	86.35	84.70	91.23	81.00	82.13	89.20	96.47	88.17	87.52
ruRoBERTa	89.67	91.63	96.68	93.62	95.60	88.83	96.17	91.10	89.83	91.64	97.20	92.40	92.86
ruGPT-small	88.53	91.57	92.94	90.33	94.30	95.33	87.63	83.20	73.17	88.82	93.93	84.30	88.67
ruGPT-medium	91.77	86.37	94.88	91.57	95.90	97.37	96.17	79.90	80.53	91.98	95.60	88.70	90.89
ruGPT-large	89.23	91.37	94.58	91.51	95.80	96.77	90.83	87.80	78.60	92.24	95.53	87.03	90.94
ruGPT-3.5-13B	94.33	95.20	97.10	96.12	97.05	98.47	98.17	94.70	87.53	96.34	97.77	95.37	95.68
SambaLingo	79.87	85.73	89.20	80.85	92.95	89.83	90.43	96.20	77.63	82.74	87.40	80.47	86.11
distil-MBERT	83.97	79.63	70.84	75.90	52.35	79.43	83.13	56.00	75.27	55.88	59.47	55.83	68.98
MBERT	88.83	84.63	78.88	80.37	86.35	87.07	82.77	52.90	66.30	61.22	59.77	52.70	73.48
XLNet-base	88.57	90.57	88.42	87.55	91.85	92.67	92.97	69.90	72.50	75.48	81.57	74.67	83.89
XLNet-large	88.80	91.03	90.52	87.73	93.15	94.37	93.07	79.10	80.67	81.30	87.70	79.77	87.27
RemBERT	51.40	54.70	48.90	49.77	32.05	51.17	62.63	45.30	51.17	49.28	52.40	52.20	50.08
MDeBERTa	52.57	43.63	47.50	36.77	75.35	41.03	37.43	40.20	43.57	41.90	44.10	53.53	46.47
mGPT-1.3B	94.37	95.97	89.64	87.69	92.15	94.20	85.13	82.50	67.80	79.94	85.43	79.53	86.20
mGPT-13B	94.53	95.53	92.08	88.75	94.35	95.33	88.50	85.60	68.27	84.46	87.70	83.03	88.18
bloom-1b7	86.10	89.70	67.86	85.55	69.75	79.10	66.87	13.10	65.20	53.78	54.67	68.87	66.71
bloom-3b	89.53	90.23	71.44	86.66	65.85	81.10	68.33	19.30	67.83	54.22	51.13	72.30	68.16
bloom-7b1	88.90	91.87	73.62	88.91	73.10	84.63	75.37	23.30	68.40	57.64	55.07	77.43	71.52
xglm-1.7B	37.70	45.03	51.72	44.26	65.70	61.57	47.23	64.10	54.63	38.20	75.27	51.93	53.11
xglm-4.5B	92.40	92.17	87.96	82.70	92.80	92.70	91.30	82.90	73.97	82.62	90.47	80.57	86.88
xglm-7.5B	92.80	93.43	88.46	83.75	93.45	93.70	91.03	90.80	74.43	83.12	90.27	82.37	88.13
Llama-7b	94.70	90.83	85.20	89.45	48.35	89.23	72.10	84.80	72.40	79.96	81.20	81.93	80.85
Llama-13b	95.83	93.50	88.50	91.23	56.00	91.53	76.97	89.00	74.00	83.08	85.50	85.63	84.23
Mistral	96.87	95.00	88.16	92.99	72.10	93.20	87.83	32.40	72.40	83.28	86.60	88.13	82.41
Human	100.0	99.33	98.80	98.53	98.0	98.67	99.33	98.00	100.0	100.0	100.0	99.33	98.62

Table 5: The average accuracy scores (%) of the 25 LMs and human baseline by phenomenon. Random baseline is 50%. The monolingual and multilingual LMs are separated by a line.