

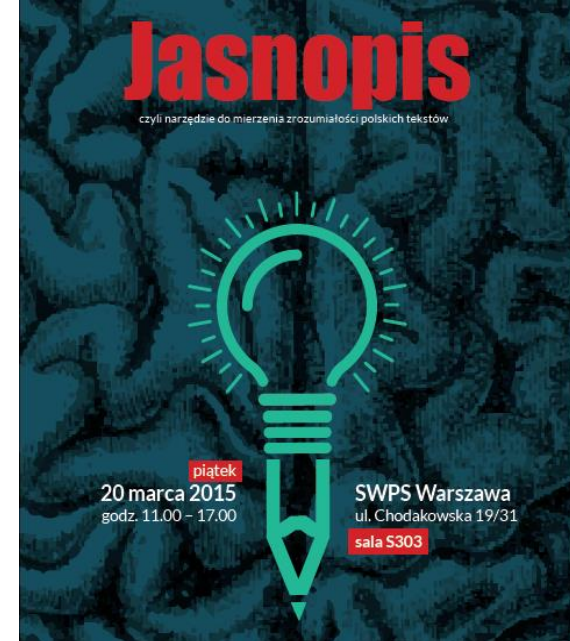


# Jak powstawał JASNOPIS, czyli narzędzie do mierzenia zrozumiałości polskich tekstów?

Wprowadzenie:

Włodzimierz Gruszczyński

Edyta Charzyńska



# FINANSOWANIE PROJEKTU

Projekt badawczy

***Mierzenie stopnia zrozumiałości polskich tekstów użytkowych (pozaliterackich)***

finansowany przez Narodowe Centrum Nauki na podstawie DEC-2011/03/B/HS2/05799 z dn. 2012-06-20.

Planowany okres realizacji: 24 mies. (08.2012 – 08.2014)

Przedłużony okres realizacji: do 15.12.2015.

Środki finansowe: 390215,00 PLN.

Jednostka koordynująca: Wydział Nauk Humanistycznych i Społecznych SWPS.

Jednostka współpracująca: Instytut Podstaw Informatyki PAN.

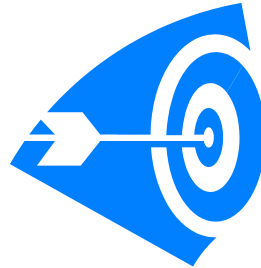


# ZESPÓŁ REALIZUJĄCY PROJEKT

- Włodzimierz Gruszczyński (kierownik, SWPS)
- Bartosz Broda (IPI PAN)
- Edyta Charzyńska (UŚ)
- Łukasz Dębowski (IPI PAN)
- Milena Hadryan (UAM)
- Bartłomiej Nitoń (IPI PAN)
- Maciej Ogrodniczuk (IPI PAN)
- Grażyna Oblas (UW)
- Laura Polkowska (UKSW)
- studentki, studenci, koleżanki, koledzy, którzy podjęli trud przeprowadzania ankiet.



# CELE PROJEKTU



- Ustalenie cech tekstu, które mają wpływ na jego zrozumiałość, z dokładniejszą czytelnością (ang. readability), przy założeniu, że badania ograniczają się w znacznym stopniu do językowych cech powierzchniowych:
  - gramatycznych (morfologia, składnia);
  - leksykalnych.
- Stworzenie narzędzia informatycznego służącego do mierzenia stopnia zrozumiałości tekstu, a w szczególności:
  - do obliczania wskaźnika czytelności badanego tekstu i ustalania tych jego cech, które odpowiadają za zbyt duży stopień trudności (np. wskazanie trudnych wyrazów, zbyt skomplikowanych zdań, zbyt długich ciągów dopełniaczowych itp.);
  - umieszczanie wyniku na skali wypracowanej w ramach projektu m.in. za pomocą metod psycholingwistycznych.

# MOTYWACJA (1)

## Istnienie zbyt trudnych tekstów w przestrzeni publicznej

- Instrukcje obsługi różnych urzędów, m.in. instrukcje bhp.
- Ulotki dołączane do farmaceutyków i kosmetyków.
- Akty prawne, np. przepisy podatkowe.
- Pisma urzędowe, zwłaszcza pisma kierowane do obywateli.
- Teksty umów bankowych, agencji turystycznych itp.
- Teksty „unijne”, np. o funduszach europejskich, o finansowaniu badań itp.
- Niektóre teksty medialne.
- Niektóre podręczniki szkolne.



# Przykład (1)

## Decyzja Miejskiego Ośrodka Pomocy Rodzinie



Miejski Ośrodek Pomocy Rodzinie podwyższa dodatek do zasiłku rodzinnego z tytułu wychowywania dziecka w rodzinie wielodzietnej na dzieci: Jadwigę B., Wandę B., i Ignacego B. o kwotę 100,- zł miesięcznie od 1 września 2007 do 31 sierpnia 2008. Podwyższony dodatek w całkowitej wysokości 180,- zł miesięcznie na każde z wymienionych dzieci będzie Pani otrzymywać w okresie od 01.09.2007 do 31.08 2008.

Uzasadnienie: Na Pani wniosek z dnia 6.09.2007 przyznano Pani dodatek do zasiłku rodzinnego z tytułu wychowania dziecka w rodzinie wielodzietnej w wysokości 80 zł miesięcznie w okresie od 01.09.2007 do 31.08.2008 na troje dzieci: Jadwigę, Wandę i Ignacego – decyzja z dn. ...

Dodatek ten podniesiono na podstawie Uchwały Rady Miasta CIII/1198/IV/2006 ... o kwotę 100 zł na trzecie i następne dzieci uprawnione do zasiłku rodzinnego. Podwyższony podatek finansowany będzie ze środków własnych gminy...

# Przykład (2)

## Fragment informacji z UE dla prasy



Panel wysokiego szczebla – powołany w celu doradzania Komisji Europejskiej w zakresie funkcjonowania wspólnego programu Ambient Assisted Living (wspólny program w zakresie życia wspieranego przez otoczenie) – rozpoczął konsultacje, w ramach których poproszono obywateli, przedsiębiorstwa oraz naukowców o podzielenie się pomysłami, jak najlepiej wykorzystać technologie informacyjne i komunikacyjne (TIK), aby pomóc osobom starszym w bardziej niezależnym życiu, a w szerszym ujęciu, aby ustanowić nowe sposoby wykorzystania rozwiązań TIK na rzecz najsłabszych członków społeczeństwa. Panelowi przewodniczy była komisarz europejska Meglena Kunewa. Konsultacje publiczne stanowią pierwszy krok w kierunku osiągnięcia celu wyznaczonego przez agendę cyfrową (zob. IP/10/581), jakim jest podwojenie do 2015 r. liczby osób starszych korzystających z udogodnień umożliwiających niezależne życie.



# Przykład (3)

## Fragment uzasadnienia wyroku sądowego



Działający z upoważnienia wnioskodawcy – Dyrektor Zarządzający – Prokurent T. P. zaproponował aby ustalenie przebiegu granicy nastąpiło po rozstrzygnięciu sprawy prowadzonej w Sądzie Okręgowym do sygn. [...] z wniosku skarżącego o wydanie nieruchomości.

Podniesiono, że w odpowiedzi na zapytania skarżącego o zaawansowanie prac, informowano skarżącego pismami, które są w aktach, że zakończenie postępowania rozgraniczeniowego warunkowane jest dostarczeniem przez Geodetę dokumentów, które w pierwszej kolejności posłużą do wydania przez Wójta Gminy opinii w trybie art. 33 ust. 2 pkt 1 ustawy z dnia 17 maja 1989 r. Prawo Geodezyjne i Kartograficzne. Po uzyskaniu tej opinii i przyjęciu dokumentów do powiatowego zasobu dokumentacji geodezyjnej i kartograficznej prowadzonego przez Starostę Powiatowego, dokumentacja będzie zdatna do wydania decyzji w przedmiocie rozgraniczenia.

# MOTYWACJA (2)

## Zwiększająca się świadomość społeczna i dążenie do zmian

- Zainicjowanie kampanii społecznej „Język urzędowy przyjazny obywatelom” przez Biuro Rzecznika Praw Obywatelskich.
- Działania takich urzędów, jak Ministerstwo Rozwoju Regionalnego, Ministerstwo Finansów, Ministerstwo Administracji i Cyfryzacji, Agencja Rozwoju Przedsiębiorczości, ZUS.
- Urzędowy wymóg sprawdzania czytelności tekstów, np.:
  - rozporządzenie Ministra Zdrowia z 26 kwietnia 2010 r. w sprawie badania czytelności ulotek dołączanych do lekarstw;
  - poradnik MEN dla rzeczoznawców oceniających podręczniki szkolne zawierający wymóg sprawdzania, czy stopień trudności tekstu odpowiada możliwościom percepcyjnym ucznia.
- Powstanie Pracowni Prostej Polszczyzny i Ruchu Prostego Języka na Uniwersytecie Wrocławskim.

# MOTYWACJA (3)

## Zapotrzebowanie na wiarygodne i wygodne narzędzie

- Potrzeba narzędzia umożliwiającego sprawdzenie stopnia trudności tekstu metodą, która
  - jest obiektywna,
  - jest szybka,
  - została zweryfikowana za pomocą badań empirycznych,
  - jest w pełni dostosowana do cech języka polskiego,
  - daje dla porównania również wyniki obliczane według innych znanych sposobów.

# CO BYŁO PRZED JASNOPISEM?

## CZYLI STAN BADAŃ I ISTNIEJĄCE NARZĘDZIA

### Na świecie

- Powstało wiele metod sprawdzających czytelność tekstów, powstały też narzędzia informatyczne służące do tego celu (głównie dla języka angielskiego).
- W niektórych krajach sprawdza się stopień zrozumiałości tekstów, które mają funkcjonować w przestrzeni publicznej, np.:
  - w USA w wielu urzędach istnieje wymóg sprawdzania tekstów pod względem czytelności przed ich przekazaniem odbiorcom;
  - w Szwecji od kilkunastu lat wdrażany jest program *klarspråk*, czyli „jasny język”, dzięki czemu udało się wyprzeć z urzędów typowy ciężki styl kancelaryjny.
- Trwają badania nad metodami upraszczania tekstów, w tym także upraszczania automatycznego.

# CO BYŁO PRZED JASNOPISEM?

## CZYLI STAN BADAŃ I ISTNIEJĄCE NARZĘDZIA

### W Polsce:

- W latach 60. Walery Pisarek stworzył i opublikował wzór, pozwalający obliczyć współczynnik trudności tekstu polskiego; wzór znany był przede wszystkim w środowiskach dziennikarzy:

$$T = \frac{\sqrt{T_s^2 + T_w^2}}{2}$$

- Stanisław Gajda podjął próby obliczania stopnia trudności tekstów naukowych (1990).
- W ostatnich latach w Pracowni Prostej Polszczyzny opracowano i udostępniono w Internecie narzędzie LOGIOS obliczające indeks mglistości dla polskiego tekstu : <http://www.logios.pl>.
- W Internecie można znaleźć programy, które obliczają indeks mglistości, zazwyczaj sprowadzają się do zaimplementowania wzoru Gunniga i wykorzystania jakiegoś dostępnego algorytmu dzielącego polskie wyrazy na sylaby, por. np.: <http://www.trzyw.pl/gunning-fog-index/>.

# PROBLEM WYJŚCIOWY

CO TO ZNACZY „(Z)ROZUMIEĆ TEKST”?

- Trudności z definicją.
- W słownikach często spotyka się błędne koło:
  - „rozumieć tekst” – to ‘pojmować treść przekazywanej w nim informacji’,
  - „pojmować” – to ‘ogarniać rozumem, rozumieć’...
- Czy można ten sam tekst różnie rozumieć?

# ROZUMIENIE TEKSTU

- Okres do lat 70.: przekonanie, że rozumienie tekstu polega wyłącznie na przetwarzaniu danych językowych.
- Procesy wnioskowania wykraczające poza czysto lingwistyczną reprezentację tekstu nie są niezbędnym elementem procesu jego rozumienia.
- Czytelnik postrzegany jako bierny, czytanie jako proces statyczny (Fodor i in., 1975; Kintsch, van Dijk, 1978).

# PRZEŁOM PSYCHOLOGICZNY

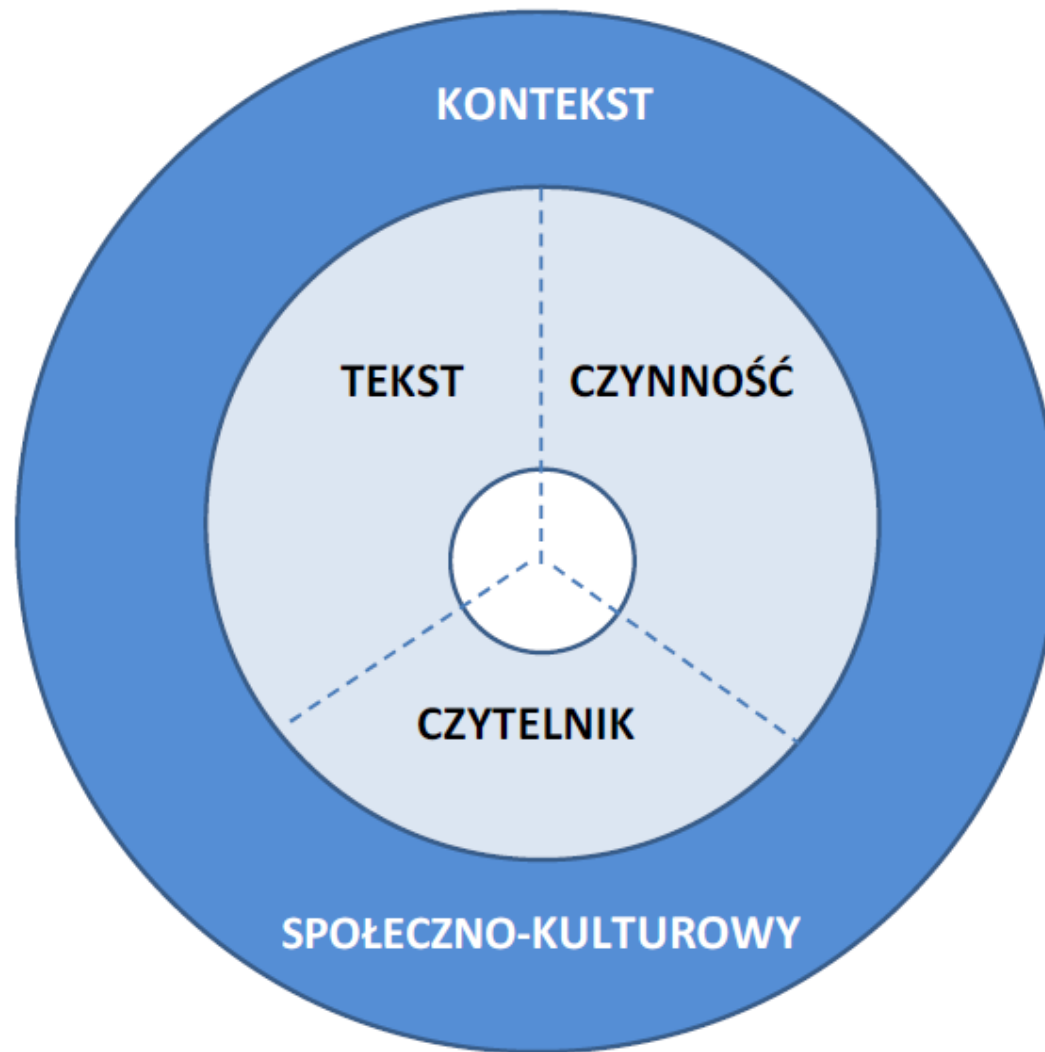
- Pełne zrozumienie tekstu wychodzi poza skonstruowanie reprezentacji czysto lingwistycznej (Johnson-Laird, 1983, Bransford, Johnson, 1972).
- Rola procesów wnioskowania (Kliś, 2000).
- Czytanie jako proces aktywny, interakcja między czytelnikiem a tekstem (NICHD, 2000; Snow, 2002).



# PRZEŁOM PSYCHOLOGICZNY

- Teoria schematów poznawczych (Anderson & Pearson, 1984; Rumelhart, 1980; Trzebiński, 2002).
- Rozumienie jako *czynny proces konstruowania reprezentacji, polegający na interpretacji odbieranych informacji zgodnie z systemem posiadanej wiedzy i na włączaniu tej informacji w ów istniejący już system* (Kurcz, 1987, s. 308).

- procesy  
GÓRA  
↑ ↓  
DÓŁ



## Model rozumienia tekstu

Rand Reading Study Group, 2002

# CZYTELNIK

- **Umiejętności poznawcze** (uwaga, pamięć, umiejętność krytycznego myślenia, wnioskowania, wizualizacje);
- **motywacja** (zainteresowanie zawartością tekstu, poczucie własnej skuteczności czytelniczej);
- **wiedza** (słownictwo, znajomość tematu, wiedza lingwistyczna, znajomość strategii rozumienia);
- **osobiste doświadczenia.**

# TEKST

- Gatunek,
- dobór słownictwa,
- styl,
- środki typograficzne,
- grafika,
- intencja autora.



# CZYNNOŚĆ I KONTEKST

## Czynność:

- **cel(e) czytania** (zewnątrzne, wewnętrzne);
- operacje służące **przetworzeniu tekstu**;
- **konsekwencje czytania** (wzrost wiedzy, rozwiązanie jakiegoś problemu oraz/lub zaangażowanie w tekst).

Składowe rozumienia pojawiają się w jakimś specyficznym **kontekście społeczno-kulturowym**.

Dziękujemy za uwagę!