

---

# DLACZEGO JEST TAK MAŁO SYNONIMÓW ?

Homonimy i synonimy w grze w nazywanie

---

Dorota Lipowska

Zakład Logiki Stosowanej  
Instytut Językoznawstwa  
UAM

- Modelowanie komputerowe staje się coraz ważniejszym narzędziem badania ewolucji języka.
- Podstawowe założenie: język to złożony system adaptacyjny, powstający na bazie lokalnych interakcji między jego użytkownikami, stopniowo komplikujący się w trakcie rozwoju zgodnego z zasadami ewolucji i samo-organizacji.
- S. Pinker i P. Bloom (1990)  
„Natural language and natural selection”

- 
- Korzyść słuchacza: zdobywa informacje.
  - Dlaczego opłaca się mówić?  
Dlaczego (stosunkowo) rzadko kłamiemy?
  - ? dobór krewniaczy
  - ? wzajemny altruizm
  - ? zdobywanie pozycji
  - ? seks
  - ? manipulowanie

- James R. Hurford (2003)  
„Why synonymy is rare: Fitness is in the speaker”
  
- model: trzy genotypy agentów konkurujące przez wiele pokoleń, aż do zdominowania populacji przez jeden z nich
  
- algorytm genetyczny preferował
  1. albo sukces komunikacyjny
  2. albo sukces interpretacyjny

1. powstaje język podobny do naturalnego: synonimy są rzadkie, występują natomiast homonimy
2. sytuacja odwrotna – nie występująca w językach naturalnych – homonimy są rzadkie, występują zaś synonimy

*Humans evolved to be well adapted as senders of messages; accurate reception of messages was less important...*

*We may be primarily speakers, and secondarily listeners.*

- Clark: wrodzona tendencja człowieka do szukania i tworzenia nowych znaczeń raczej niż akceptowania tego samego znaczenia dla różnych form.
- Markman: dzieci zakładają, że znaczenia żadnych dwóch słów nie nakładają się.
- Wexler – formalne zasady akwizycji języka: Zasada Jednoznaczności powstrzymuje dziecko przed przyswajaniem więcej niż jednej formy dla danego znaczenia.

- Homonimy są w języku naturalnym znacznie częstsze niż synonimy, choć synonimia nie wpływa na efektywność komunikacji, podczas gdy homonimia może ją pogarszać!
- Ta asymetria wydaje nam się ważną cechą, typową dla języków naturalnych, która może być wykorzystana jako test dla różnych komputerowych modeli rozwoju języka.

- Gra w nazywanie
  - ➔ dwóch agentów
  - ➔ n obiektów
- Homonimia i synonimia
  - ➔ homonimia z upływem czasu nie znika („dynamiczna pułapka”)
  - ➔ rola synonimii wyraźnie maleje (cecha przemijająca)
- **Asymetria między homonimią a synonimią może być więc wytłumaczona w ramach dość prostego modelu gry w nazywanie, bez odwoływania się do argumentu ewolucyjnego Hurforda (słuchacz czerpie większe korzyści z konwersacji niż mówca).**



- **Luc Steels (1995) – model naming game:**  
grupa komunikujących się ze sobą agentów próbuje ustalić wspólne słownictwo dla pewnej liczby obiektów (zwykle 1).
- Tylko wymiana kulturowa (w ramach jednej generacji).
- Stan lingwistycznej zgodności (*linguistic coherence*).
- **Iterated Learning Model** – interakcje międzypokoleniowe.

- 
- Dwa agenty: mówca i słuchacz (na zmianę)
  - Mówca wybiera obiekt, następnie reprezentujące go słowo i komunikuje je słuchaczowi.
  - Słuchacz odgaduje znaczenie słowa.
  - Sprawdzenie – sukces lub porażka determinują modyfikacje słownika

- Każdy agent z każdym z  $n$  obiektów wiąże pewien zestaw odpowiadających mu słów (maksymalnie  $l$ ).
- Każdemu słowu przypisana jest waga  $w$ .
- Słowa to liczby całkowite z przedziału  $\langle 1, r \rangle$  (parametr  $r$  określa wielkość dostępnej przestrzeni werbalnej).

## ■ Mówca

- losowo wybiera obiekt
- z odpowiadającej mu listy słów wybiera komunikowane słowo  $x$   
(losowanie ruletkowe względem wagi słów)

## ■ Słuchacz

- dla każdej listy ( $k=1,n$ ) oblicza miarę jej podobieństwa do usłyszanego słowa  $x$ :

$$s^k(x) = \frac{1}{\sum_i w_i} \sum_i \frac{w_i}{\varepsilon + |x_i - x|}$$

$10^{-5} \leq \varepsilon \leq 10^{-1}$  - zapewnia skończoność miary  $s^k$

- ➔ traktując obliczone miary jako wagi, wybiera w losowaniu ruletkowym listę (obiekt)

## ■ Modyfikacja list

- ➔ jeśli lista słuchacza ma ten sam numer co mówcy (sukces), to obaj agenci zwiększają wagi komunikowanego słowa (jeśli słuchacz nie ma go na danej liście, to dodaje je z wagą jednostkową)
- ➔ w przeciwnym przypadku (porażka) mówca zmniejsza wagę słowa, a słuchacz (na liście o numerze wybranym przez mówcę) zwiększa jego wagę lub w przypadku jego braku – dodaje je
- ➔ *reinforcement learning approach*

## ■ Wpływ szumu

- ➔ z prawdopodobieństwem  $p$  komunikowane jest słowo

$$x_c = x + \eta$$

$-a \leq \eta \leq a$  ( $a$  - amplituda szumu,  
 $\eta$  - liczba losowa)

- ➔ z prawdopodobieństwem  $1-p$  komunikowane jest słowo  $x$ .

$n=4, l=3, k=2$

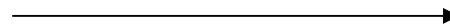
**Mówca**

1244
5678
6879
7890
5458
3609
7823
5678
9751
1134
9974
5001

**Słuchacz**

1244
5667
1221
6658
7892
1012
7823
1230
9751
1244
9342
2381

7890



**SUKCES**

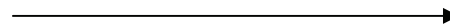
## Mówca

1244
5678
6879
7890
5458
3609
7823
5678
9751
1134
9974
5001

## Słuchacz

1244
5667
1221
6658
7892
1012
7823
1230
9751
1244
9342
2381

7890



# PORAŻKA



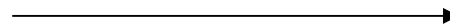
## Mówca

1244
5678
6879
7890
5458
3609
7823
5678
9751
1134
9974
5001

## Słuchacz

1244
5667
1221
6658
7892
1012
7823
1230
9751
1244
9342
2381

7891

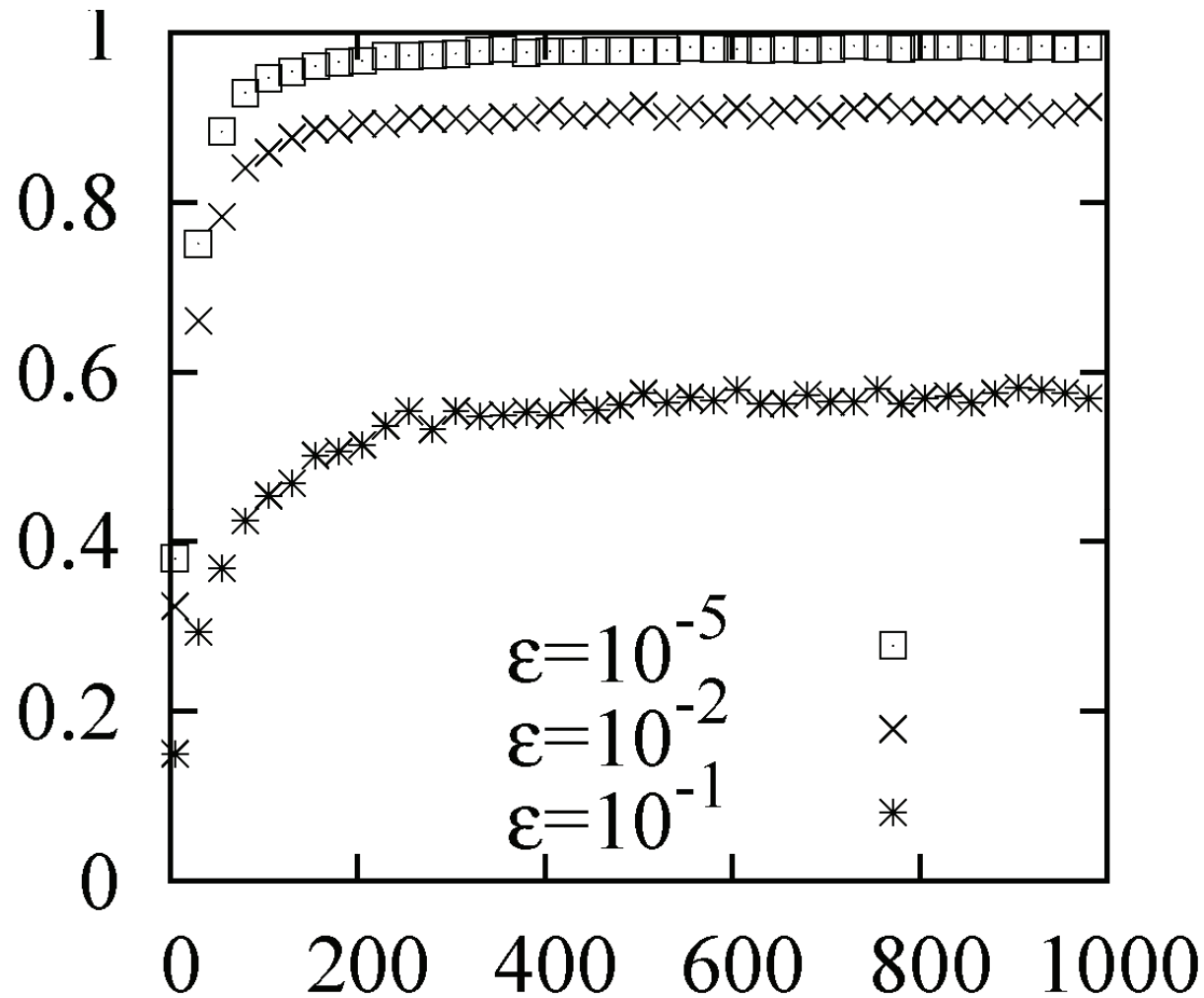


**SUKCES W OBECNOŚCI SZUMU**

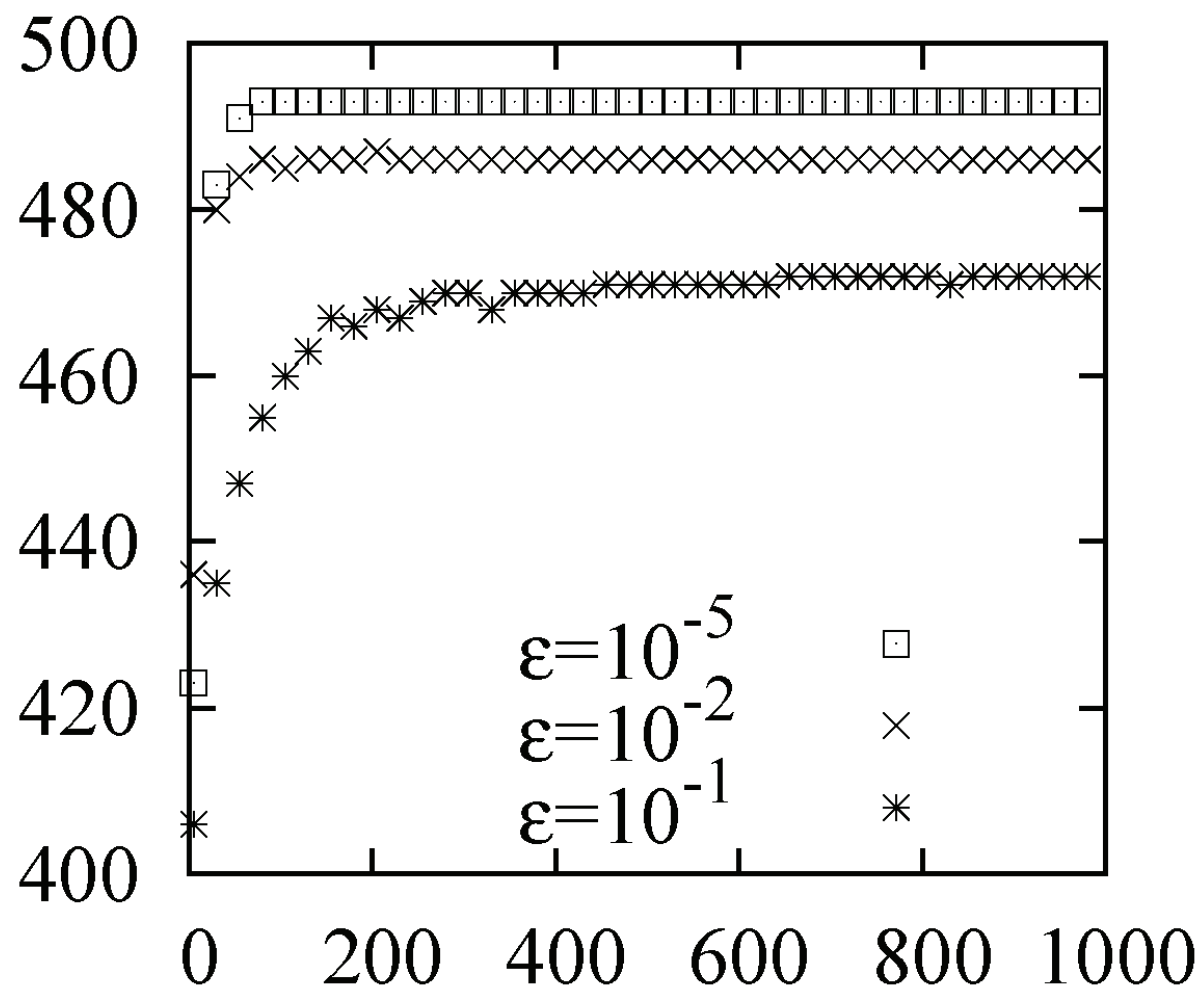
parametr	opis i wartości
n	liczba obiektów $(100 \leq n \leq 1000)$
l	maksymalna liczba słów odpowiadających obiektowi $(5 \leq l \leq 20)$
r	słowa to liczby naturalne nie przekraczające r $(500 \leq r \leq 10000)$
$\varepsilon$	zapewnia skończoność miary podobieństwa $(10^{-5} \leq \varepsilon \leq 10^{-1})$
p, a	parametry opisujące szum $(0 \leq p \leq 0.05, \quad 0 \leq a \leq 10)$

# OBLICZENIA NUMERYCZNE

- Konfiguracja początkowa: każdy agent ma na każdej liście jedno słowo wybrane losowo i o wadze jednostkowej.
- Jednostka czasu –  $2n$  prób komunikacji.
- Agenci korelują swoje listy, osiągając dość duży sukces komunikacyjny.

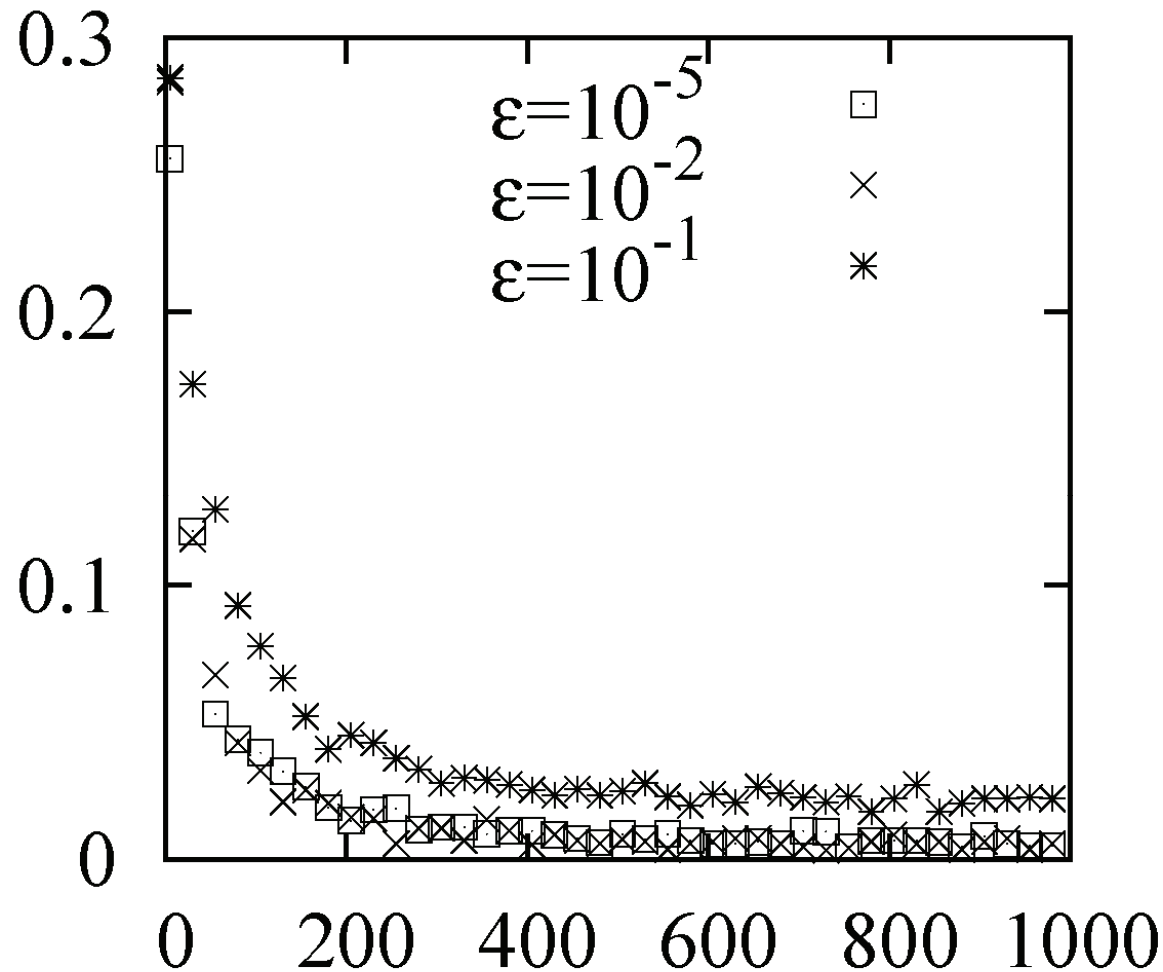


**Stosunek liczby sukcesów do liczby wszystkich prób komunikacji (w funkcji czasu) dla  $n=500$ ,  $l=10$ ,  $r=1000$**



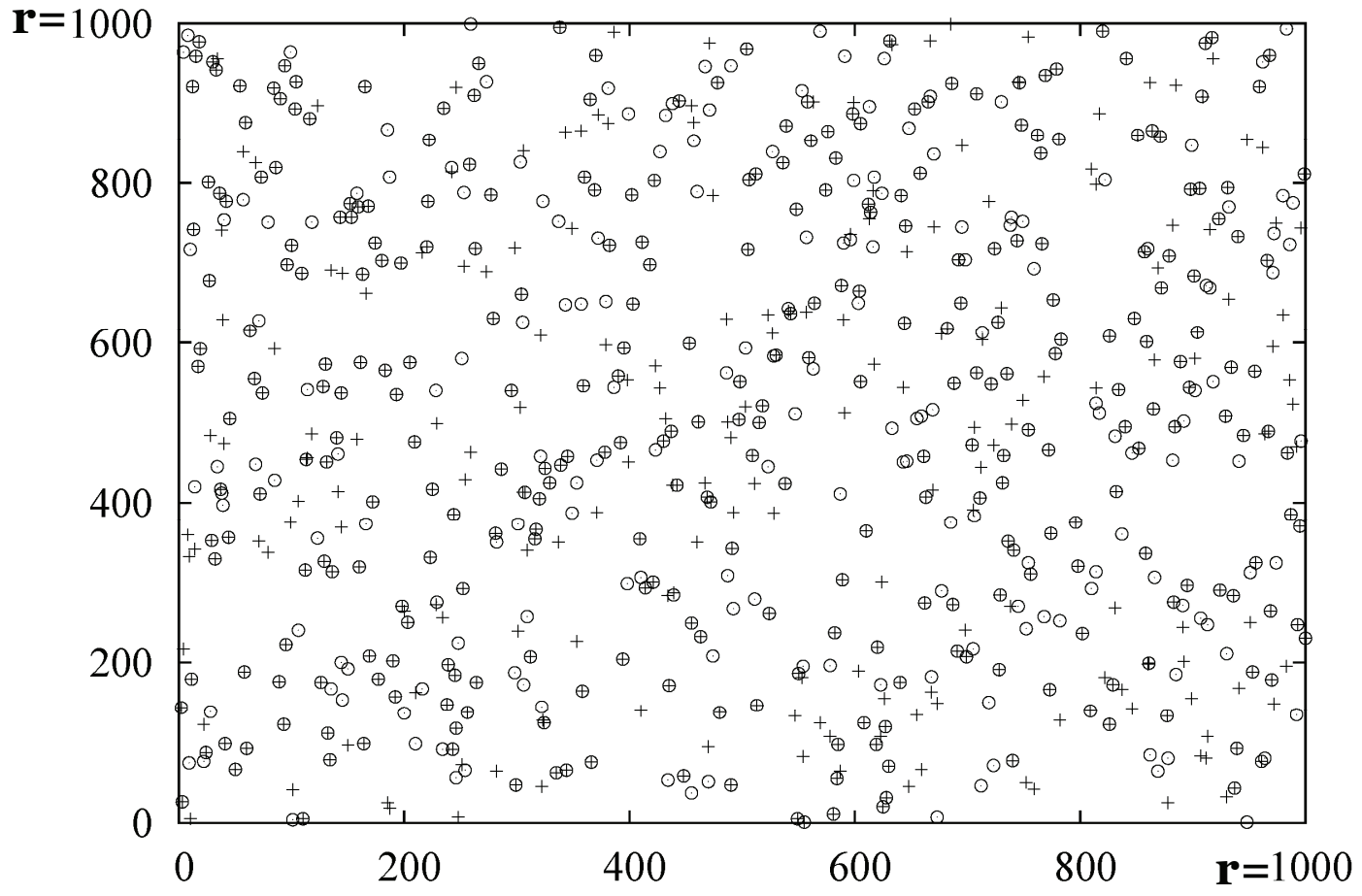
**Liczba różnych „najcięższych” słów (w funkcji czasu)  
dla  $n=500$ ,  $l=10$ ,  $r=1000$**





**Stosunek liczby wypowiedzi przy użyciu słów drugich co do wagi do liczby wszystkich prób komunikacji (w funkcji czasu) dla  $n=500$ ,  $l=10$ ,  $r=1000$**





**Dystrybucja słów o największej wadze (oś pozioma) i drugich co do wagi (oś pionowa) dla obu agentów (○ i +) dla  $n=500$ ,  $t=1000$ ,  $r=1000$**



---

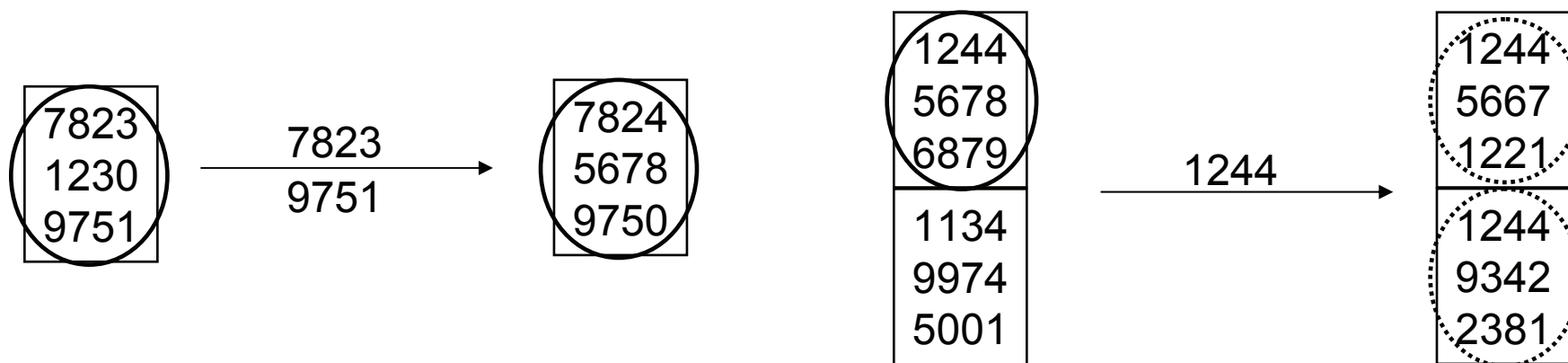
# HOMONIMIA i SYNONIMIA

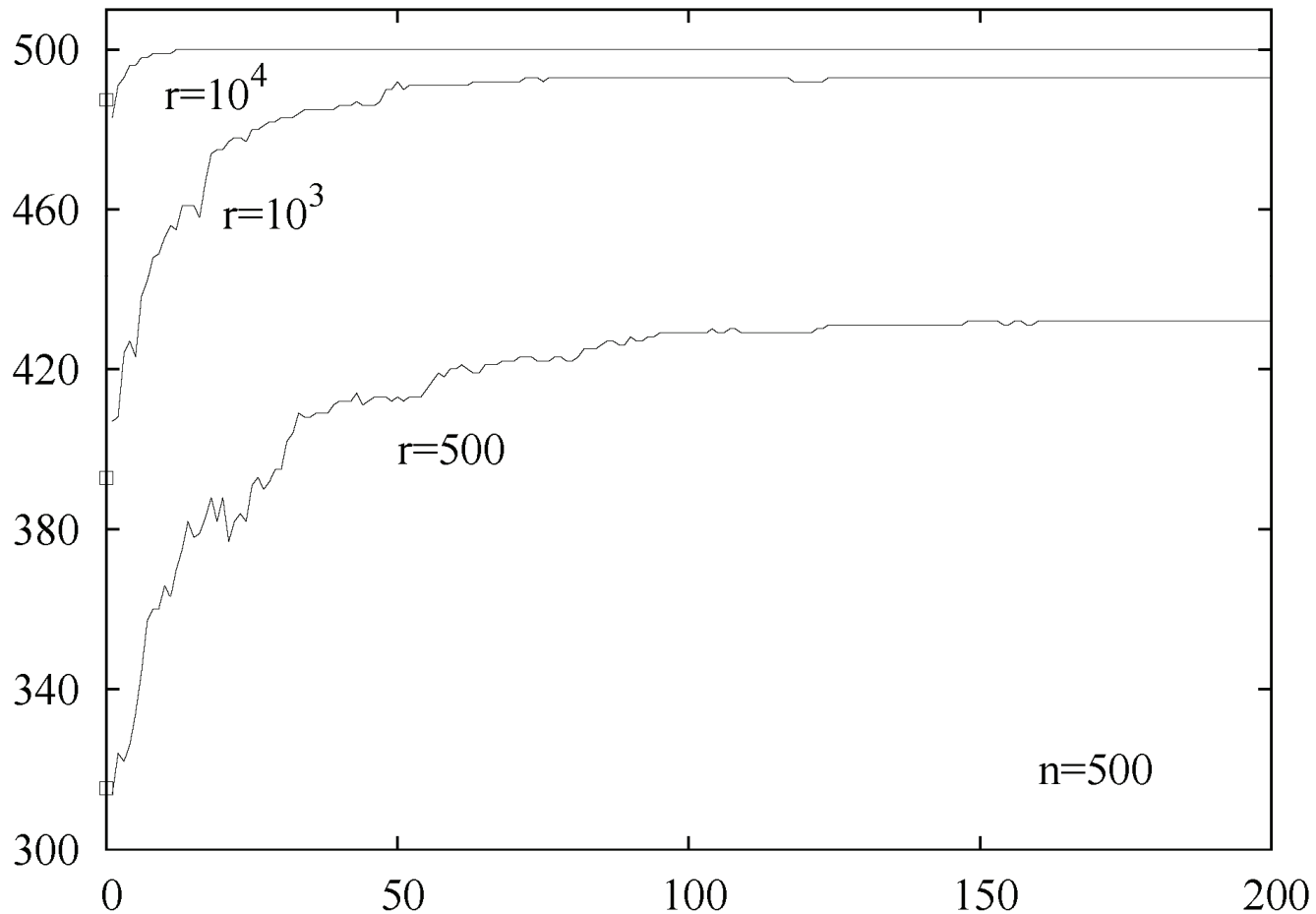
- Homonim – słowo, które może być kojarzone z więcej niż jednym obiektem
- Synonim – słowo przypisane obiektowi, z którym skojarzone jest więcej niż jedno słowo
- Elementy probabilistyczne modelu



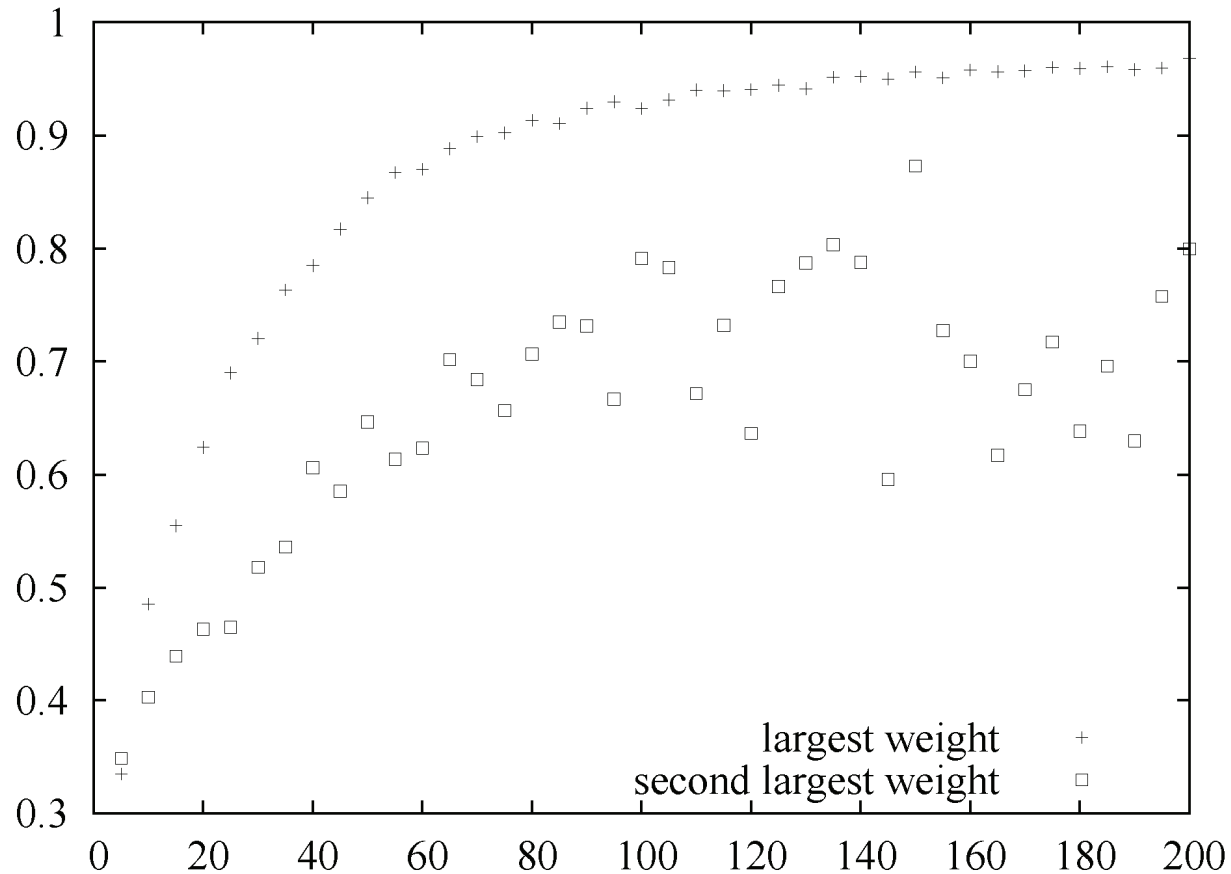
- **HOMONIM** – słowo, które ze stosunkowo dużym prawdopodobieństwem może oznaczać różne obiekty.
- Taka sytuacja występuje zwykle, gdy słowo wybrane przez mówcę występuje na więcej niż jednej liście słuchacza jako „najcięższe”.
- Miarą homonimii języka jest więc liczba różnych „najcięższych” słów: im jest mniejsza, tym częstsze są homonimy.

- **SYNONIMY** – słowa, które z dość dużym prawdopodobieństwem mogą odnosić się do tego samego obiektu.
- Taka sytuacja występuje zwykle, gdy oba agenty na tej samej liście mają te same (lub bardzo zbliżone) „najcięższe” słowa oraz słowa drugie co do wagi.





**Liczba różnych słów o największej wadze dla  $n=500$ ,  $l=10$ ,  $\varepsilon=10^{-5}$  w funkcji czasu ( $\square$  - wartości dla słów wybranych losowo)    **▶▶****



**Stosunek liczby sukcesów do liczby wszystkich prób komunikacji (w funkcji czasu) przy użyciu słów „najcięższych” i drugich co do wagi**



- Homonimia chociaż rzadka jednak nie znika (stała cecha języka).
- Po okresie początkowym częstotliwość wypowiedzi homonimicznych utrzymuje się na stałym poziomie.
- Częstotliwość wypowiedzi synonimicznych maleje z czasem.


- Przewidywania modelu dla języków naturalnych (synonimia w nich będzie zjawiskiem rzadkim) są zgodne z obserwacjami.
- Objaśnienie tego zjawiska zaproponowane przez Hurforda: wywołuje je asymetria między korzyściami ewolucyjnymi mówcy i słuchacza.
- Nasz model – znacznie prostszy, bez efektów ewolucyjnych, rozwój języka w ramach jednej generacji (tylko mechanizmy kulturowe).

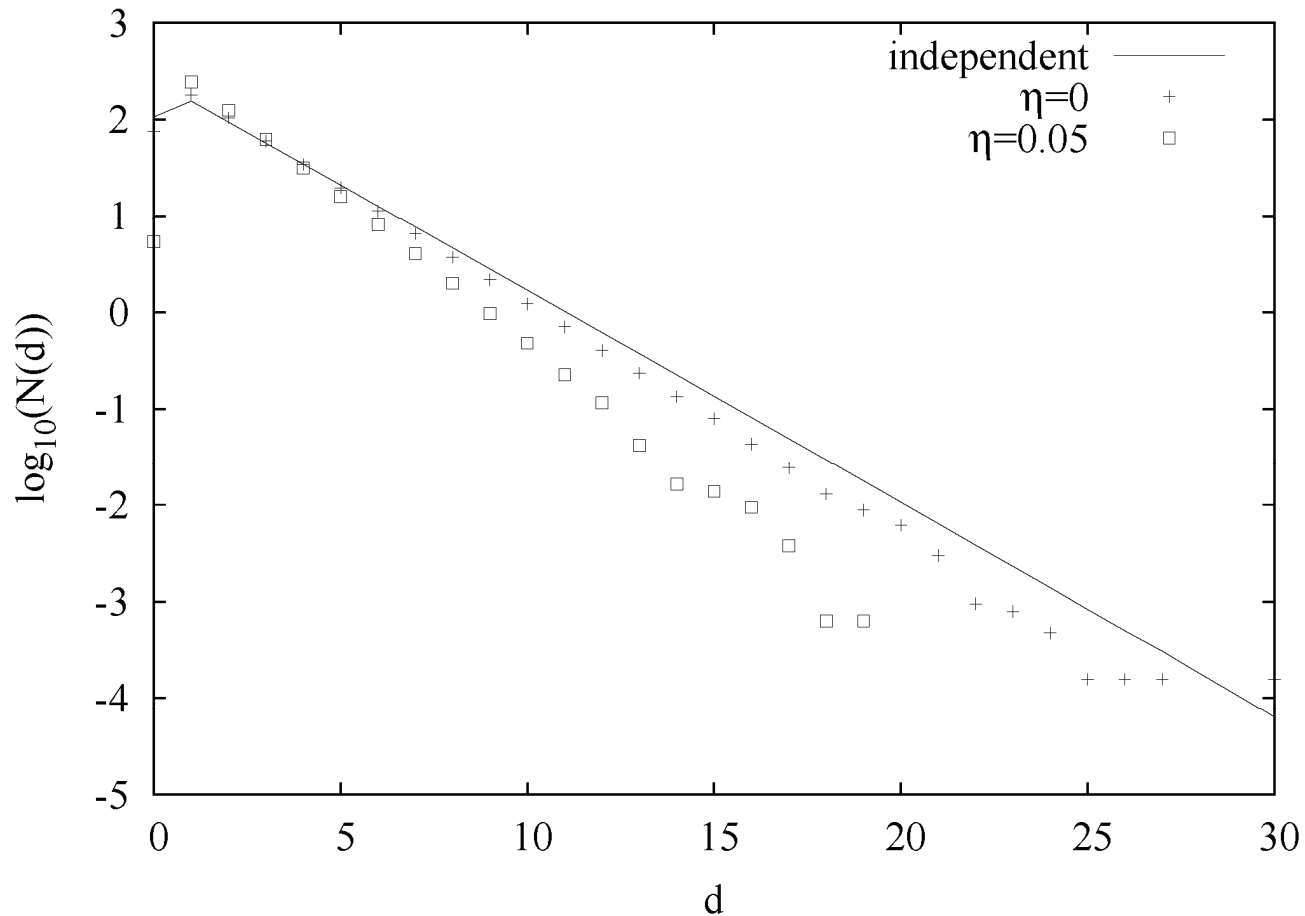
# SZUM a DYSTRYBUCJA

z prawdopodobieństwem  $p$  komunikowane jest słowo

$$x_c = x + \eta$$

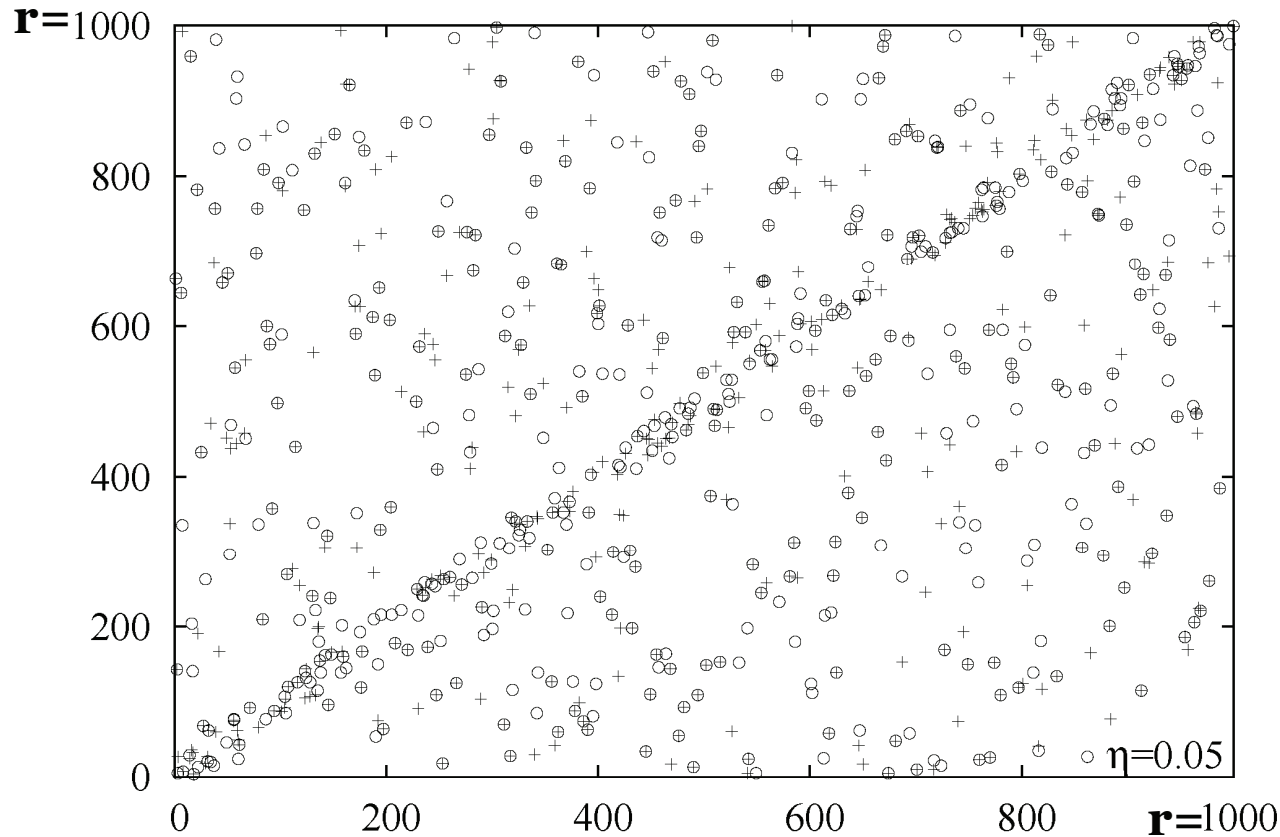
$-a \leq \eta \leq a$  ( $a$  - amplituda szumu,  $\eta$  - liczba losowa)

- Już przy  $p=0$  następuje redystrybucja słów „najcięższych” redukująca homonimię. 
- Szum istotnie wpływa na dalszy wzrost tej redystrybucji słów.



**N(d) – średnia liczba odległości d między sąsiednimi słowami „najcięższymi” (n=500, r=1000, l=10, ε=10<sup>-5</sup>)**





**Dystrybucja słów o największej wadze (oś pozioma) i drugich co do wagi (oś pionowa) dla obu agentów ( $\circ$  i  $+$ ) dla  $n=500$ ,  $t=1000$ ,  $r=1000$**



- Prawdopodobnie szum odegrał ważną rolę w procesie ewolucji języka:
  - ➔ Wpłynął korzystnie na redystrybucję słów w obrębie dostępnej przestrzeni werbalnej.
  - ➔ Zredukował liczbę homonimów.
  - ➔ Zredukował liczbę synonimów.

---

# W przyszłości ...

- Model w wersji wieloagentowej ?
- Model z ewolucją agentów ?
- Model z interakcją danego języka z innymi ?

---

DZIĘKUJĘ



