

The semantics of poetry: A distributional reading

by Aurélie Herbelot
University of Cambridge, UK

Евгения Бельская

Лингвистика – Обычный язык – Поэзия

- Поэтический язык \leftrightarrow общие правила синтаксиса и семантики
- Особая «поэтическая семантика»
- Значение поэзии \leftrightarrow значение обычного (непоэтич.) языка

*Особая задача лингвистики: работать с языком, когда он используется для творческих целей

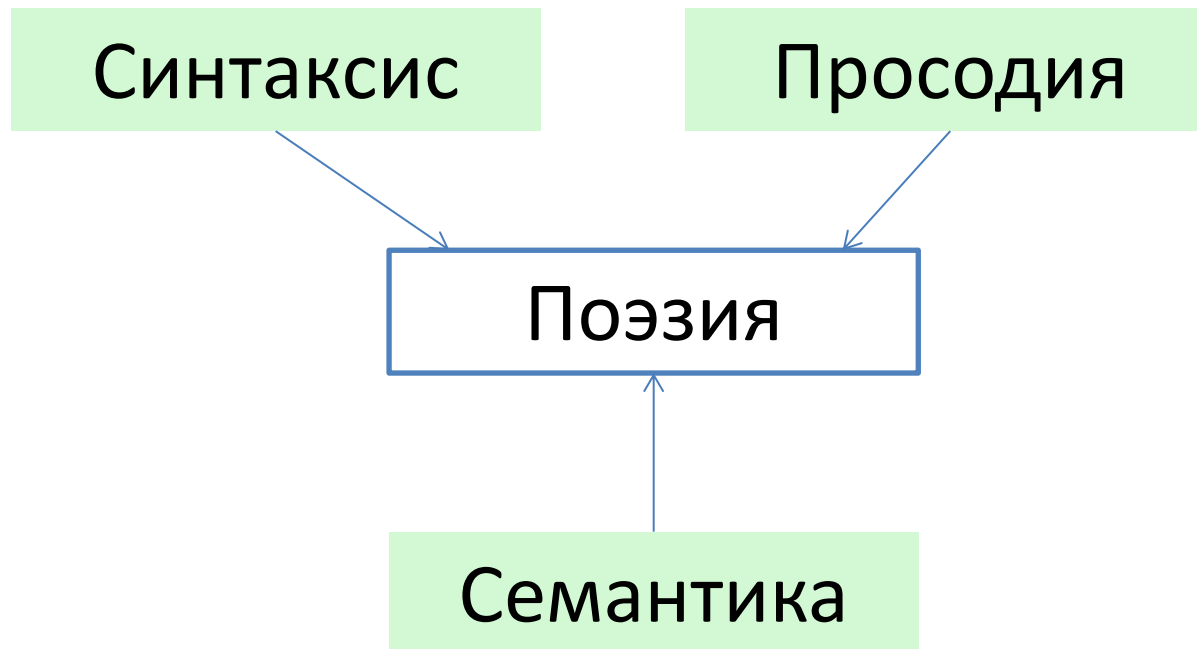
Гипотеза

- Поэзия – это форма языка, которая может быть проанализирована лингвистически
- Поэтическая семантика использует структуру обычного (непоэтического) языка для создания смыслов

Значение слова



Теория: лингвистический анализ поэтического языка



Теория: дистрибутивная семантика

- Контекст определяет значение слова (Л.Витгенштейн)
- Слова что-либо обозначают в мире (проблема нескольких контекстов)

Word

Star

Extension



Intention

Morning Star

Evening Star

Практическая часть. Задачи

1. Создать вычислительную модель значения слова по корпусу непоэтического языка
2. Применить ее к поэтическому тексту
3. Выявить, каким образом поэзия основывается на повседневном дискурсе

Материал

- 8 поэтических текстов 1881-2008 гг.
- 1 статья из Википедии
- 1 автоматически составленный текст

Title	Author	Year	Excerpt
Day that I have loved	Rupert Brooke	1911	Tenderly, day that I have loved, I close your eyes,/ And smooth your quiet brow, and fold your thin dead hands.
Argument over, Amounting	Clark Coolidge	1990	In edges, in barriers the tonal light of t/ the one thing removed overemphasizes tonally/ and you could hurry it, and it vanish and plan
Valentine	Carol Ann Duffy	1993	Not a red rose or a satin heart./ I give you an onion./ It is a moon wrapped in brown paper.
Five A.M.	Allen Ginsberg	1996	Elan that lifts me above the clouds/ into pure space, timeless, yea eternal/ Breath transmuted into words/ Transmuted back to breath
At Issue III	Karen Mac Cormack	2001	Putting shape into getting without perfect in a culture that doesn't think, pumps up, the/ two traits go at the face of rate themselves
Ithaca, Winter.	Avery Slater	2008	Creaking, skeleton bells of trees/ dissolve in a quilt of pale flurries.
If I Told Him, A Completed Portrait of Picasso	Gertrude Stein	1924	If I told him would he like it. Would he like it if I told him./ Would he like it would Napoleon would Napoleon would would he like it.
In The Gold Room— A Harmony	Oscar Wilde	1881	Her ivory hands on the ivory keys/ Strayed in a fitful fantasy,/ Like the silver gleam when the poplar trees/ Rustle their pale-leaves listlessly
'The Language Poets'	Wikipedia	?	The Language poets (or L = A = N = G = U = A = G = E poets, after the magazine of that name) are an avant garde group or tendency in United States poetry that emerged in the late 1960s and early 1970s.
Psychologist. Strong.	Random text	-	Tabard, battersea, wolf, coma, acas. hutchinson cap'n. suet. ellesmere. proportionality/ mince. outside, morey folk, cum, willoughby, belligerent, dimension

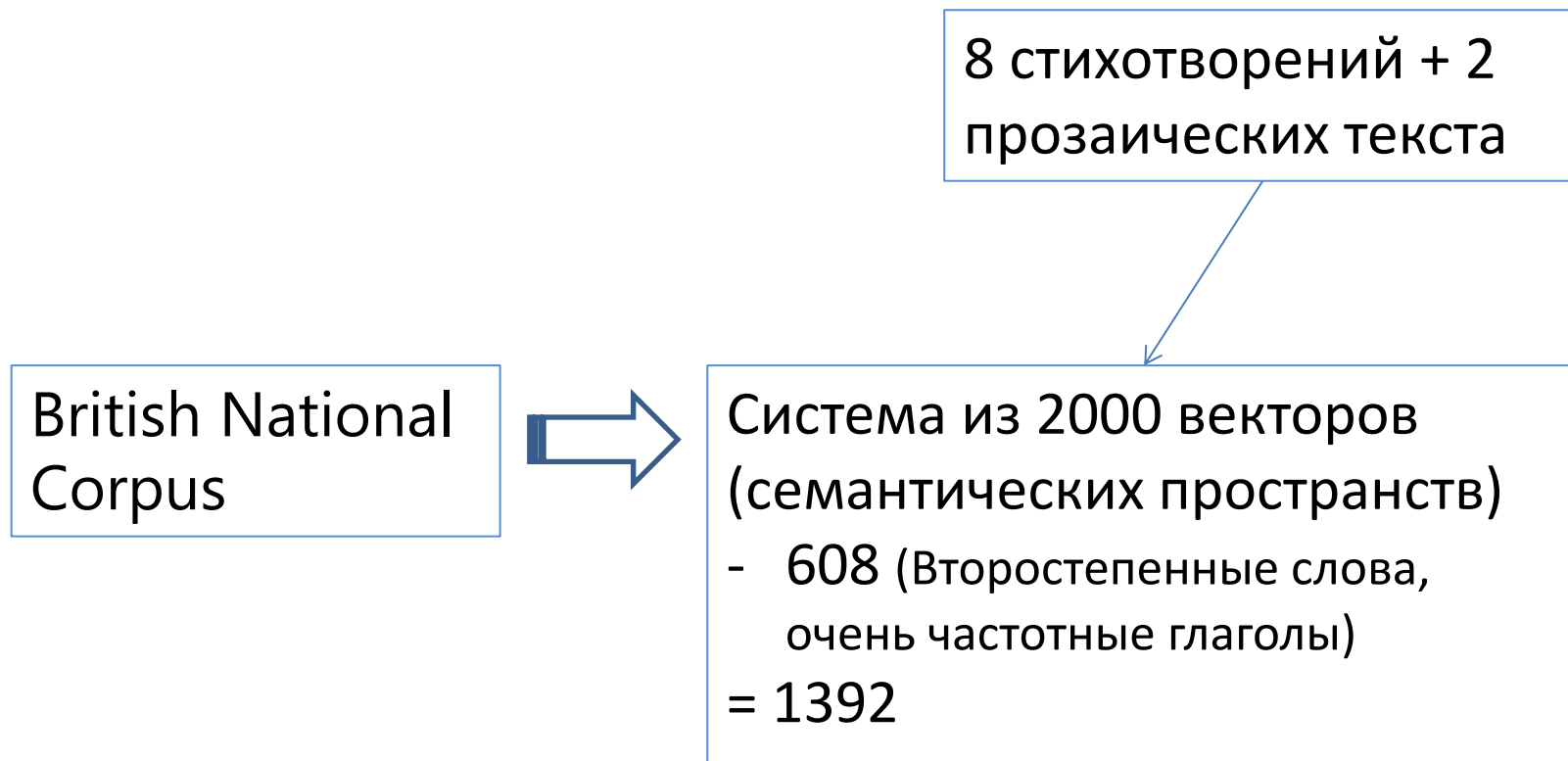
Семантическая связность в поэзии

- Связность нескольких слов – это значение их попарного сходства
- “the reaches of turning aside” (Coolidge, 1900)

$$\textit{MeanSimScore}(w_i \dots w_n)$$

$$= \textit{mean}\{\textit{Sim}(w_i, w_j), ij \in 1 \dots n, i < j\}$$

Семантическая связность в поэзии



Семантическая связность vs. сложность текста

Table 2 Difficulty scores for each text in sample

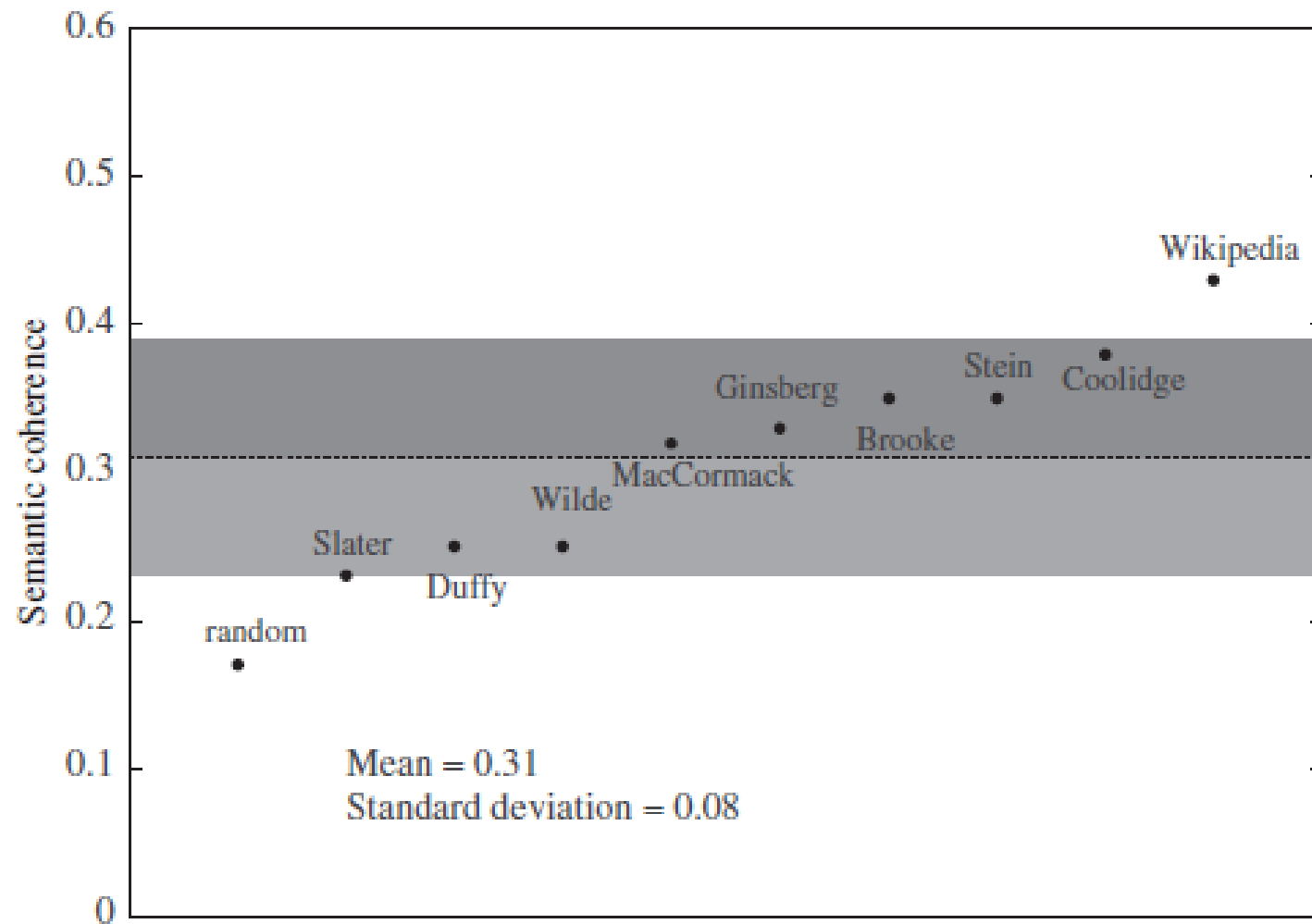
	Author	Annotator 1	Annotator 2	Average
Random	5	5	5	5.00
Mac Cormack	5	5	5	5.00
Coolidge	4	5	5	4.67
Ginsberg	5	4	3	4.00
Stein	5	3	3	3.67
Slater	2	3	4	3.00
Brooke	2	4	3	3.00
Wilde	1	1	2	1.33
Duffy	1	1	2	1.33
Wikipedia	1	1	1	1.00

Семантическая связность vs. сложность текста

Table 3 Semantic coherence for each text in sample

Poem	Average coherence	Subjective difficulty scores (average)
Random	0.17	5.00
Slater	0.23	3.00
Duffy	0.25	1.33
Wilde	0.25	1.33
Mac Cormack	0.32	5.00
Ginsberg	0.33	4.00
Brooke	0.35	3.00
Stein	0.35	3.67
Coolidge	0.38	4.67
Wikipedia	0.43	1.00
MEAN	0.31	
STDV	0.08	

Семантическая связность vs. сложность текста



Семантическая связность vs. сложность текста

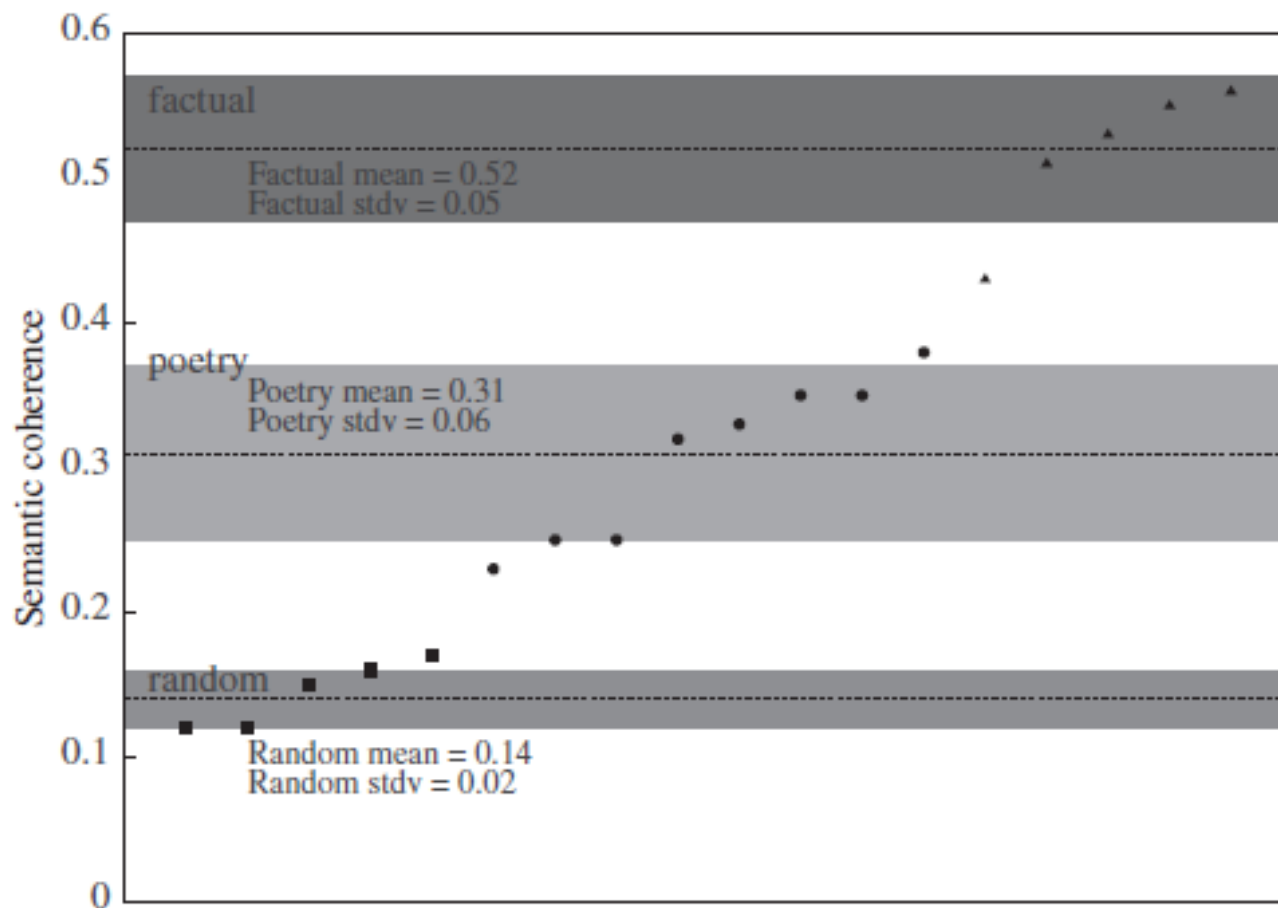


Fig. 4 Coherence range for poetry versus random and factual texts

Семантическая связность vs. сложность текста

Table 4 Similarities and shared contexts for the fragment *play arrived, how large in prompting*

Word pair	Similarity	Shared contexts
play_N–arrive_V	0.49	{ticket_N scene_N studio_N hall_N} {australia_N africa_N} {wednesday_N minute_N regularly_A}
play_N–large_A	0.44	{audience_N hall_N} area_N flat_A dominate_V
play_N–prompt_V	0.43	{tv_N audience_N performance_N success_N} {write_V version_N scene_N story_N} {violence_N dominate_V} {move_N run_N} united_A rain_N
arrive_V–large_A	0.56	{ship_N station_N island_N} {crowd_N hall_N}
arrive_V–prompt_V	0.48	{police_N troop_N army_N warning_N} {finally_A eventually_A june_N march_N weekend_N} {flight_N visit_N} {paris_N germany_N} news_N scene_N miss_N william_N couple_N
large_A prompt_V	0.46	{audience_N gather_V} {firm_N organization_N europe_N} {coal_N plastic_N fruit_N} complex_A volume_N dominate_V

- Семантическая прозрачность – это не фиксированное свойство словосочетания, но, скорее, состояние ума читателя

Выводы

- Через дистрибутивное понимание значения возможно показать отношение между непоэтическим языком и экстраординарным поэзии

Дистрибутивная модель

- показывает разницу между текстом, написанным человеком, и автоматическим текстом
- Стабильный уровень семантической ассоциативности среди всех стихотворений, несмотря на их сложность