

# Языковые сети: структура, функции и эволюция

Докладчик: Каменская Анна, НИУ-ВШЭ

Статья: Language Networks: their structure, function and evolution (Ricard V. Sole, Bernat Corominas Murtra, Sergi Valverde, Luc Steels<sup>3</sup>, )

# План презентации

- Языковые универсалии
- Типы языковых сетей
- Сети смежности
- Синтаксические сети
- Семантические сети
- Статистические универсалии языковых сетей
- Рост и эволюция языковых сетей
- Выводы

# Языковые универсалии

## Классификация универсалий



Абсолютные

Статистические

Импликативные  
(Сложные)

## Причины возникновения универсалий

1. Строение человеческого организма и его мозга
2. Функция языка – коммуникация
3. Взаимосвязь между языками
4. Язык как сложная сеть

# Язык как сложная сеть

## Основные понятия и измерения в теории сложных сетей

$W = \{ W_i \}$ , где  $i = 1, \dots, N$  – множество слов

$\{ W_i, W_j \}$  – пара слов, их связь задает отношение "слово – слово"

Количество связей – **степень узла**

$L$  – общее число связей

$\langle k \rangle = \frac{2L}{N}$  – **средняя степень**

**Длина пути:** среднее минимальное расстояние между любыми двумя вершинами

**Коэффициент кластеризации:** вероятность того, что у двух связанных вершин есть третья, связанная с ними обеими

**Случайные графы:** каждая пара вершин соединяется с какой-то вероятностью (например, граф Эрдёша-Реньи); низкий коэффициент кластеризации и короткая длина пути

**Структура «тесный мир»:** высокий коэффициент кластеризации, но короткая длина пути

**Распределения степеней связанности узлов:** функция распределения степеней узлов  $P(k)$ , которая определяется как вероятность того, что узел  $i$  имеет степень  $k_i = k$

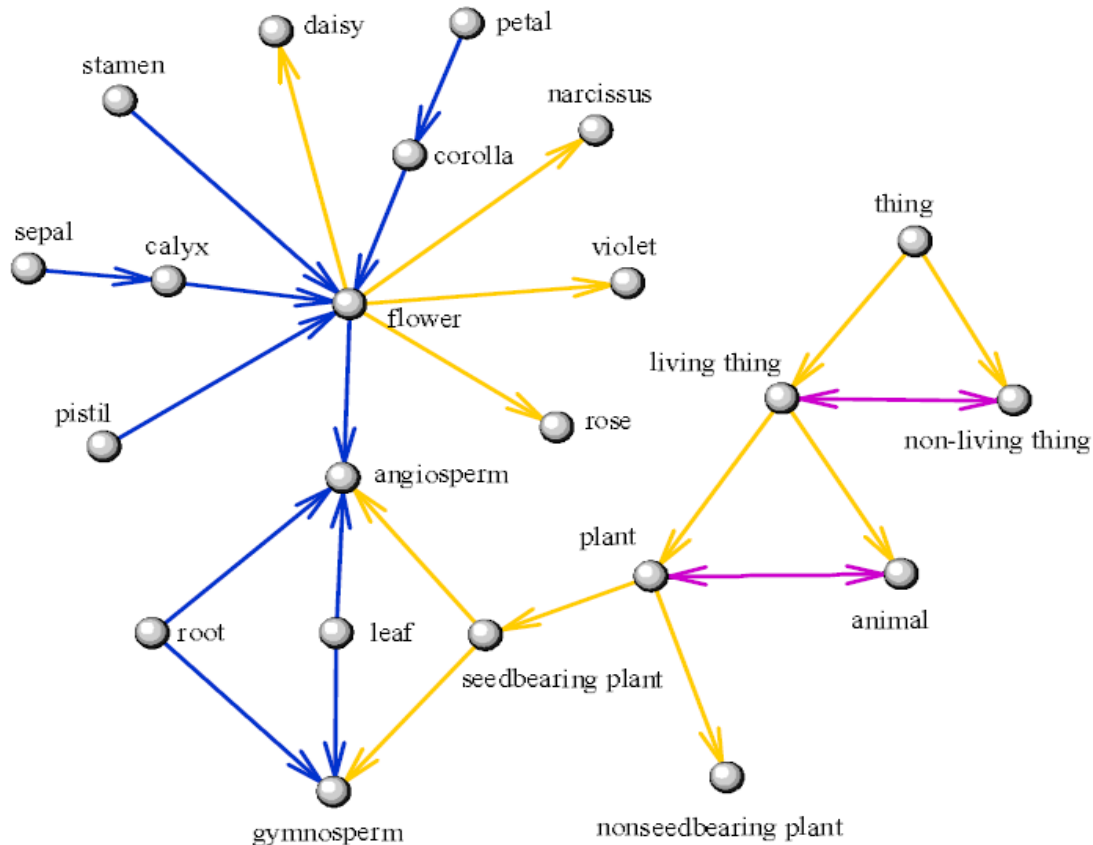
# Типы языковых сетей

- **По уровням языка:**
  - Синтаксический
  - Лексический (содержит семантический)
  - Морфемный
  - Фонемный
- **По типу языковых обществ** (Principles of Linguistic change. Volume II: Social Factors, Labov, W. (2000) ; Network Topology and self-consistency in Language Games, Corominas, Bernat and Sole Ricard (2005))
- **По семиотическим подходам** (Sharp Transition towards Shared Vocabularies in Multi-Agent Systems, Baronchelli, A., Felici, M., Caglioti, E., Loreto, V., and Steels, L. (2005); Collective learning and semiotic dynamics, Steels, L. and Kaplan, F. (1999))





# Семантические сети



Желтые ребра – isa-relation (гиперонимия)

Синие ребра – меронимия (отношение часть-целое)

Фиолетовые ребра – антонимия

# Статистические универсалии языковых сетей

|                        | Co-occurrence networks          | Syntactic Networks              | Semantic Networks              |
|------------------------|---------------------------------|---------------------------------|--------------------------------|
| Size                   | $N \sim 10^3 - 10^6$            | $N \sim 10^3 - 10^4$            | $N \sim 10^4 - 10^5$           |
| Average degree         | $\langle k \rangle \sim 4 - 8$  | $\langle k \rangle \sim 5 - 10$ | $\langle k \rangle \sim 2 - 4$ |
| Link relation          | Precedence                      | Syntactic dependence            | Word-word association          |
| Functional meaning     | Sentence production             | Grammar architecture            | Psycholinguistic web           |
| Path length            | $d \sim 3 - 4$                  | $d \sim 3.5$                    | $d \sim 3 - 7$                 |
| Clustering             | $C/C_{rand} \sim 10^3$          | $C/C_{rand} \sim 10^3$          | $C/C_{rand} \sim 10^2$         |
| Link distribution      | SF, $\gamma \sim 2.2 - 2.4$     | SF, $\gamma \sim 2.2$           | SF, $\gamma \sim 3$            |
| Hubs                   | Words with low semantic content | functional words                | Polysemous words               |
| Effect of hub deletion | Loss of optimal navigation      | Articulation loss               | Loss of conceptual plasticity  |

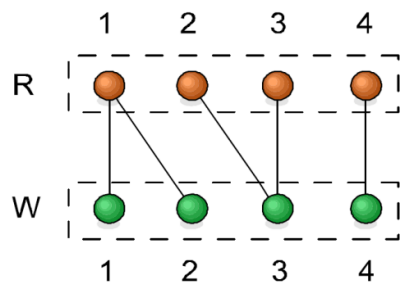
# РОСТ И ЭВОЛЮЦИЯ ЯЗЫКОВЫХ СЕТЕЙ

## Принципы формирования языковых сетей:

- Когнитивный потенциал
- Окружение

## Влияние принципов формирования языковых сетей на статистические свойства их роста:

- Успешные коммуникации (например, Naming Games)



R – значения

W – слова

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

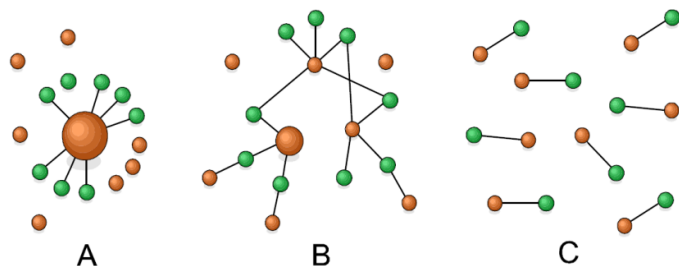
- Принцип минимизации усилий

$E_s$  – усилия говорящего

$E_h$  – усилия слушающего

$\lambda$  – параметр, где  $0 < \lambda < 1$

$$\Omega(\lambda) = \lambda E_s + (1 - \lambda) E_h$$



# Выводы

- Языковые сети содержат статистические свойства схожие с признаками других сложных сетей:
  - Определение новых статистических свойств языка
  - Применение законов, использующихся в сложных динамических системах к языку

Спасибо за  
внимание!