

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ФОНОЛОГИЧЕСКИХ СЕТЕЙ

Бажанов Д.С., Прогулова Т.Б.

Международный университет природы, общества и человека «Дубна», Россия,
141980, Дубна, ул. Университетская, 19, тел. 2190723, E-mail: bazhds@gmail.com

Данная работа посвящена математическому моделированию фонологической сети русского языка и сравнению ее топологии с топологиями фонологических сетей английского, испанского, мандаринского, гавайского и баскского языков [1]. Целью исследования является определение топологической близости различных языков.

В фонологических сетях слова языка представляются вершинами, и две вершины связываются ребром, если слова звучат похоже, различаясь точно в одну фонему [1, 2]. Нами разработан алгоритм, позволяющий строить фонологические сети русского языка на основе учения о фонеме Московской фонологической школы. С помощью данного алгоритма была построена сеть из ~150 000 слов.

В отличие от сетей слов, фонологические сети являются несвязанными [3]. Мы показываем, что топология полученной сети корректно описывается q -экспоненциальным распределением, и определяем параметр неэкстенсивности. Были рассчитаны такие характеристики как коэффициент кластеризации и длина среднего кратчайшего пути, что позволило установить мало-мировой характер сети (небольшую длину среднего кратчайшего пути при высоком коэффициенте кластеризации). Анализ зависимости средней степени ближайших соседей вершины от ее степени показывает, что сеть является ассортативной. Также исследована структура сообществ в сети.

Используя результаты работы [1] мы проанализировали топологии фонологических сетей английского, испанского, мандаринского, гавайского и баскского языков, определили значения параметра неэкстенсивности и сравнили меры сложности этих систем, что позволило определить степень топологической близости этих языков.

Мы также показываем, что построение фонологических сетей русского языка на основе других определений (других фонологических школ) не приводит к существенным структурным изменениям.

Литература

1. *Arbesman S., Strogatz S., Vitevitch M.* The Structure of Phonological Networks Across Multiple Languages // arXiv:0903.0872v1., 2009.
2. *Vitevitch M.* What can graph theory tell us about word learning and lexical retrieval? // *Journal of Speech, Language, and Hearing Research Vol. 51, 2008. P. 408–422.*
3. *Gadjiev B.R., Progulova T.B.* Comparative analysis of collaboration networks. // *Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP Conference Proceedings 1305, 2010.*