Inside or Nearby: Capturing Phonological Similarity in the Mental Lexicon with Network Analysis of Embeddings and Neighbors

Vladislav Zubov[1, 2, 3]

[1] HSE University; [2] Institute for Linguistics Studies RAS, [3] Russian National Library

vladzubov21@gmail.com

Background. Phonological similarity affects the process of spoken word recognition, as words that sound alike can compete during lexical access. This competition arises from word onset overlap (cohorts), phonological proximity (neighbours), and the embedding of one word within another (lexical embeddings). Recently, network analysis has been increasingly employed to model the mental lexicon, enabling the construction of multilayered networks that reflect the complex interconnections of word-related information across different linguistic levels, including the phonological level. In such networks, phonological similarity is typically represented as links between word-nodes differing by a single phoneme (see, for example, [1, 2, 3, 4], among others). In this study, we present a novel approach to describing phonological similarity using network analysis based on a different principle – lexical embeddings. We assume that the topological structure of a phonological embedding network (PEN) will differ from that of a phonological neighbourhood network (PNN) and will offer new insights into the role of phonological similarity in speech perception.

Method. Two corpora were used for the analysis: the open-access Britfone corpus, which contains 16,205 phonetic transcriptions of English words [5], and the English Lexicon Project (ELP) database, which provides word frequency data and reaction times in a lexical decision task [6]. The two datasets were merged, stress markers were removed (as is standard practice in neighbourhood networks), and words with identical transcriptions were combined. Reaction times were averaged, and frequencies summed. A total of 15,405 unique transcriptions were analysed. We constructed two networks: a directed PEN, where a link was established between nodes if one word was fully embedded within another (e.g., *cat* < *catalogue*), and a PNN, where a link was formed between words differing by a single phoneme (e.g., *cat* – *bat* – *at*). Key network metrics were computed for comparison. For each node, we calculated the number of embeddings and neighbours, as well as various network metrics (closeness, clustering coefficient, PageRank). We then built a multilevel linear regression model with reaction time from the ELP as the dependent variable, and network metrics, word frequency, and word length as predictors.

Results. The two networks differed in terms of key topological metrics. Although PEN was less dense, it was more connected and less fragmented than PNN. This reflects a more hierarchical structure of phonological similarity, characterised by numerous short, central words into which longer words are embedded (see Table 1). Comparisons of several linear models showed that the best-performing model included network metrics from both networks. The results of the linear regression indicated that an increase in the number of embedded words facilitated reaction times ($\beta = -0.05$, $SE = 0.01$, $t = -4.29$, $p < 0.001$), whereas an increase in the number of neighbours slowed reaction times ($\beta = 0.20$, $SE = 0.01$, $t = 14.37$, $p < 0.001$) – see the full model output in Figure 1.

Discussion. We demonstrated that network metrics from PEN provide novel insights into word recognition. The observed facilitatory effect of more embeddings on reaction time contradicts earlier assumptions regarding embeddings as competitors [7]. Meanwhile, the inhibitory effect of a greater number of neighbours aligns with previous findings [8].

Table 1. Network metrics for the Phonological Embedding Network and Phonological Neighbourhood Network

Metric	PEN	PNN
Nodes	15 049	11 502
Edges	48 250	36 565
Density	0.0002	0.0006
Components	101	1654
Mean total degree	6.412	6.358
Mean distance	2.912	6.399
Transitivity (CC)	0.018	0.316
Assortativity degree	0.020	0.710
Centralization degree	0.068	0.003

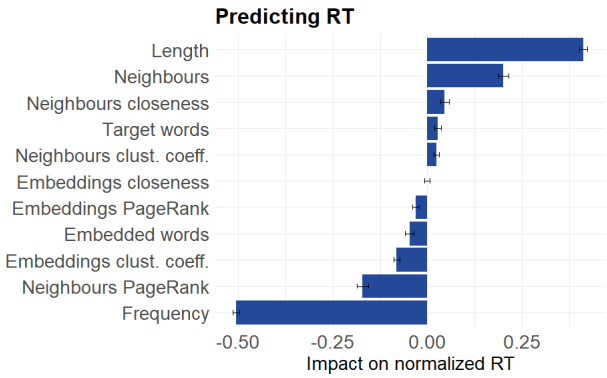

Figure 1. Standardised Beta Coefficients from the Linear Model Predicting Lexical Decision RTs

References

[1] Vitevitch, M. S. (2008). What Can Graph Theory Tell Us About Word Learning and Lexical Retrieval? Journal of Speech, Language, and Hearing Research, 51(2), 408–422. https://doi.org/10.1044/1092-4388(2008/030)

[2] Siew, C. S. Q. (2013). Community structure in the phonological network. Frontiers in Psychology, 4. https://doi.org/10.3389/fpsyg.2013.00553

[3] Levy, O., Kenett, Y. N., Oxenberg, O., Castro, N., De Deyne, S., Vitevitch, M. S., & Havlin, S. (2021). Unveiling the nature of interaction between semantics and phonology in lexical access based on multilayer networks. Scientific Reports, 11(1), 14479. https://doi.org/10.1038/s41598-021-93925-y

[4] Stella, M., Citraro, S., Rossetti, G., Marinazzo, D., Kenett, Y. N., & Vitevitch, M. S. (2024). Cognitive modelling of concepts in the mental lexicon with multilayer networks: Insights, advancements, and future challenges. Psychonomic Bulletin & Review, 31(5), 1981–2004. https://doi.org/10.3758/s13423-024-02473-9

[5] Llarena, J. (2017). Britfone: British English (RP/Standard Southern British) pronunciation dictionary. GitHub. https://github.com/JoseLlarena/Britfone

[6] Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. Behavior Research Methods, 39(3), 445–459. https://doi.org/10.3758/BF03193014

[7] McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. Language and Cognitive Processes, 10(3–4), 309–331. https://doi.org/10.1080/01690969508407098

[8] Vitevitch, M. S., & Luce, P. A. (2016). Phonological Neighborhood Effects in Spoken Word Perception and Production. Annual Review of Linguistics, 2(1), 75–94. https://doi.org/10.1146/annurev-linguistics-030514-124832