What governs a language's lexicon? Determining the organizing principles of phonological neighbourhood networks

Rory Turnbull and Sharon Peperkamp

Abstract The lexicons of natural language can be characterized as a network of words, where each word is linked to phonologically similar words. These networks are called phonological neighbourhood networks (PNNs). In this paper, we investigate the extent to which observed properties of these networks are mathematical consequences of the definition of PNNs, consequences of linguistic restrictions on what possible words can sound like (phonotactics), or consequences of deeper cognitive constraints that govern lexical development. To test this question, we generate random lexicons, with a variety of methods, and derive PNNs from these lexicons. These PNNs are then compared to a real network. We conclude that most observed characteristics of PNNs are either intrinsic to the definition of PNNs, or are phonotactic effects. However, there are some properties—such as extreme assortativity by degree—which may reflect true cognitive organizing principles.

1 Introduction

In natural languages, sentences are composed of words, which are in turn composed of strings of symbols referred to as *phonemes*, which represent the smallest units of sound that can be used to distinguish words from each other. Many psycholinguistic theories of spoken word recognition and infant language acquisition rely on a concept of the phonological similarity of words, termed *neighbourhood*, which is defined in terms of the phonemic structure of words. Two words are neighbours of each other if they differ by the deletion, addition, or substitution of one and only one segment—that is, an edit distance of one. For example, neighbours of *plan* include

Rory Turnbull (e-mail: rory.turnbull@ens.fr)✉ · Sharon Peperkamp (e-mail: sharon.peperkamp@ens.fr)
Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS)

Rory Turnbull · Sharon Peperkamp
Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University, 29 rue d'Ulm, 75005 Paris, France

H. Cherifi et al. (eds.), *Complex Networks & Their Applications V*,
Studies in Computational Intelligence 693,
DOI 10.1007/978-3-319-50901-3_7

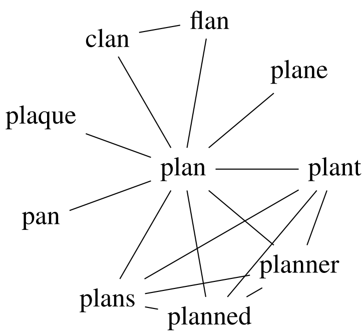

Fig. 1 Example phonological neighbourhood network centred around the English word *plan*. Note that it is the sound of a word, not the spelling, which determines the phonological neighbours. Note further that some neighbours of a word are neighbours of each other.

pan (deletion of /l/), *plant* (addition of /t/), and *clan* (substitution of /k/ for /p/). See Figure 1 for a visual example.[1] The neighbourhood relation is symmetric, intransitive and anti-reflexive.

For a given lexicon, then, it is possible to construct a complex network to model phonological neighbourhood relations throughout the language. Phonological neighbourhood networks (PNNs) have been used to study aspects of lexical organization in several languages [1, 16, 20]. In this paper, we explore the extent to which these complex network analyses can provide insight into the psychological organization of human language.

Vitevitch [20] first proposed the use of PNNs to study the phonological aspects of lexicons. In a PNN, every word in the lexicon is a vertex in a graph, and two vertices are linked by an edge if a neighbourhood relation obtains between the two words. This process yields an undirected, unweighted graph, ideal for examination with the tools of complex network analysis.

Early work on PNNs, in a variety of languages, has demonstrated that these networks have distinct properties which differ in important ways from other complex networks studied in the literature [1, 20]. For example, while most complex networks typically have a giant component which contains around 80–90% of the vertices, the observed values for PNNs fall between 10% and 65% [1, 16]. PNNs were also found to be remarkably robust to vertex removal, with the average shortest path length remaining the same when up to 5% of vertices were removed. Notably, this effect held regardless of whether vertex removal was at random or in order of degree [1]. Despite these differences from other networks, the high clustering coefficients established that PNNs exhibit small world properties.

However, these statistics and examinations rely on comparing the observed networks to random networks [14]. While this approach is reasonable for many kinds of complex networks, it is not an appropriate comparison for PNNs. Unlike other networks, where vertices exist independently of each other and edges can be made

[1] Note that neighbourhood is defined based on the pronunciation of a word, not the spelling. For instance, while the spelling of the words *knee* and *neat* are quite different, the pronunciations are very similar. The addition or deletion of the /t/ sound will transform *knee* into *neat* and vice versa. Therefore, these words are neighbours. On the other hand, the words *tough* and *though* have very similar spellings, but their pronunciations—/tʌf/ and /ðoʊ/ respectively—are very different. These words are not neighbours.

or unmade (for example, friendships made or broken, shipping routes established or abandoned), in a PNN the edges (neighbourhood relations) are intrinsic to the definition of the vertices themselves (the phonological structure of the words). That is, because edges exist between two vertices if and only if the two words are phonological neighbours, there are certain graphs which are not possible PNNs.

One such graph is shown in the left of Fig. 2. Here, each vertex is connected to every other vertex, with two exceptions: vertices 1 and 3 are not connected, and vertices 2 and 4 are not connected. It not possible for this graph to have its vertices labelled such that the shortest path from vertex-to-vertex is equal to the edit distance (Hamming distance) of the vertex labels [9]. In other words, this graph cannot represent neighbourhood relations between words. On the other hand, the graph on the right of Fig. 2 is plausibly a PNN, with the mapping 0=*cant*, 1=*can*, 2=*cat*, 3=*cab*, 4=*cap*.[2] Note that the graphs in Fig. 2 both have the same number of vertices and edges, but the left one could not be a PNN while the right one could be.

The difference between these graphs is that the graph on the right is *addressable*, that is, there exists a vertex labelling schema which satisfies the neighbourhood relation, while the graph on the left is non-addressable [2]. Addressable graphs have also been termed ℓ_1-graphs, as it can be shown that addressable graphs are isometrically embeddable into a hypercube [6, 17]. Since the distances along the edges of a hypercube fall under the definition of an ℓ_1 metric, it follows that these graphs are isometrically embeddable into an ℓ_1 metric space [7]. The recognition of such graphs can be solved in polynomial time [8, 11].

For these reasons, random graphs are inappropriate as comparison cases when considering PNNs. Currently, it is not easy to tell if results obtained are generalizable results about language and lexical organization, or if they are simply consequences of the structure of an addressable graph [10]. It has further been noted that the statistics of PNNs are very sensitive to the distribution of word lengths within a lexicon and the number of phonemes in the language's symbol set [16, 19]. For example, given *n* phonemes, the number of possible words is an exponential term of

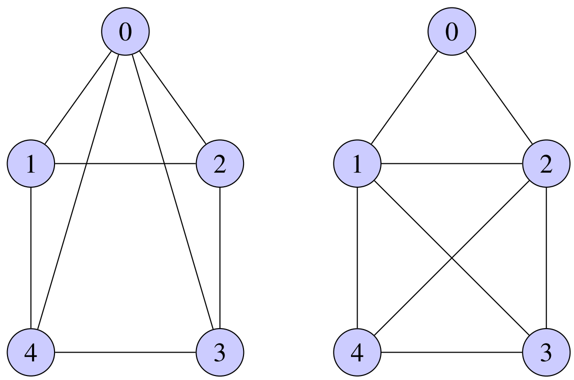

Fig. 2 Two graphs, both with the same number of vertices and edges. The graph on the left is non-addressable. This graph could not represent a PNN. The graph on the right is addressable. This graph could be represent a PNN. Consider the mapping 0=*cant*, 1=*can*, 2=*cat*, 3=*cab*, 4=*cap*. Each vertex is connected to its phonological neighbours.

[2] Other possible mappings include (0=*slow*, 1=*low*, 2=*sew*, 3=*go*, 4=*show*); (0=*lamp*, 1=*lamb*, 2=*lap*, 3=*lab*, 4=*lad*); (0=*gasp*, 1=*gas*, 2=*gap*, 3=*gag*, 4=*gash*); (0=*iode*, 1=*eyed*, 2=*ode*, 3=*aid*, 4=*add*) and so on.

n, while the number of possible neighbourhood connections is a linear term of n [19]. This fact has consequences for how cross-linguistic comparisons are carried out, as languages differ in the sizes of their lexicons and their number of phonemes [16]. These difficulties make the use of complex network analysis in the study of PNNs a complex undertaking.

In this study, we generate random lexicons, rather than random graphs. PNNs are derived from these random lexicons, guaranteeing that the resulting graphs are addressable. These simulated PNNs can be compared to real PNNs. In broad terms, there are two possible outcomes to this investigation:

1. The simulated PNNs are indistinguishable from a real PNN.
2. The simulated PNNs differ from a real PNN.

In the case of (1), we can conclude that alleged properties of the human language faculty relating to lexical organization [1] are simply consequences of the mathematical structure of PNNs. In this regard, the results could shed light on the hypercube-embeddable graphs, but not on language.

In the case of (2), we can conclude that any areas of difference between the simulated PNN and the real PNN are due to some organizing principle or cognitive constraint operating on language. For example, to ensure efficient communication, the lexicon may be organized to avoid having words which sound very similar [12].

2 Method

To address the question of which properties of PNNs are simply due to their definition and which are due to linguistic principles, we generated random lexicons, derived PNNs from these lexicons, and compared the properties of these PNNs to the PNN of English. The PNN of English we used was derived from the Hoosier Mental Lexicon [15], a dictionary of American English with phonological transcriptions of 19,320 words, after homophone removal. We refer to this lexicon and PNN as the 'real English lexicon' and 'real English PNN' to distinguish it from the simulated (random) lexicons and PNNs that we generated.

2.1 Random lexicons

Each random lexicon had the same size and mean word length (6.35 phonemes), and used the same inventory of phonemes, as the real English lexicon. Five groups of random lexicons were generated, differing in the extent to which they approximate the real English lexicon: uniform random lexicons; Zipfian random lexicons; scrambled random lexicons; bigram random lexicons; and trigram random lexicons. Each group consisted of 200 random lexicons.

The simplest group was the **uniform random lexicons**, which were created by randomly sampling from the phoneme inventory in a unform manner. Word length was sampled from a Poisson distribution (with $\lambda = 6.35$). In these lexicons, while the overall properties of the lexicon (number and length of words) was the same as that of the real English lexicon, the content of the words resemble what one would obtain from random typing.

Zipfian random lexicons were created in the same manner, except that the sampling from the phoneme inventory was not uniform. Instead, phonemes were frequency ranked according to a Zipf distribution. That is, given N phonemes, the probability of phoneme ϕ_k, where $k \in \{1, \ldots, N\}$ is given as

$$p(\phi_k) = \frac{k^{-1}}{\sum_{n=1}^{N} n^{-1}}.$$

Phoneme distributions in natural languages are approximately Zipfian [21]; these lexicons therefore approximate more closely the structure of English than the uniform random lexicons.

The **scrambled random lexicons** began with the real English lexicon and scrambled the order of the phonemes within each word. This scrambling disrupts the neighbourhood structure of the words, while preserving the overall phoneme frequencies exactly.

Of these three groups, the uniform random group approximates the average word length of English; the Zipfian group the average word length and average phoneme frequency; and the scrambled group matches word length and phoneme frequencies exactly. An important difference between these groups and the real English lexicon is that of *phonotactics*—higher-level generalizations about the combinatoric possibilities of phonemes. The classical example is that neither *blick* nor *bnick* are actual English words, but the former could be a word, while the latter could not. This is due to a restriction in what consonant clusters English permits at the beginning of syllables.[3]

Due to the lack of phonotactics in the randomly generated lexicons, any differences between them and the real English lexicon could either be due to organizing principles of lexical storage, or simply a consequence of the fact that phonotactics restrict the possible words that can appear in a lexicon. To test for this possibility, the bigram and trigram random lexicon groups were generated.

These random lexicons were generated by creating n-gram models of English phoneme distributions, where $n = 2$ for the **bigram random lexicons** and $n = 3$ for the **trigram random lexicons**. In these models, the probability of a given phoneme is conditioned on the probability of the preceding $n - 1$ phonemes. (Kneser-Ney discounting was applied to smooth the probability space for unobserved forms.) In this way, the model is able to account for basic distributional facts of English phonotactics—for example, vowels and consonants tend to alternate; the consonant cluster ‘thl’ (as in *decathlon*) is rare, but the consonant cluster ‘str’ (as in *string*) is common; and so on. Using this model, a lexicon the same size as the real English lexicon was generated. Due to the fact that the n-gram models encodes the probability of individual phonemes, and the ‘end-of-word’ character, these generated lexicons

[3] Note that in some languages, like Russian, both *blick* and *bnick* are possible words, while in others, like Japanese, neither are possible words.

approximate the real English lexicon in terms of phoneme frequencies and mean word length.

The bigram model yields English-like words, but there are exceptions, for example, /#nd/, where # represents the beginning of a word. There are no English words that begin with /nd/.[4] This situation arises due to the fact that the model can only 'see' two phonemes at a time. The sequence /#n/ (that is, the beginning of a word, followed by /n/) is a frequent bigram sequence, and so it has relatively high probability; likewise, the sequence /nd/ is frequent and also has a relatively high probability, and so therefore there is a chance that the model will output sequences like /#nd/. The trigram model, on the other hand, is able to see three phonemes at a time, notes that /#nd/ is not attested in the original lexicon, and accordingly assigns this sequence an extremely low probability. Thus, the trigram model is more English-like than the bigram model. Still, phonotactics are considerably more complex than phoneme-level n-gram probabilities, and the trigram model still produces words which sound quite un-English-like. The use of complex phonotactic generators to create 'English-like' simulated lexicons can help alleviate this problem [12], but such an investigation is beyond the scope of the current study.

To summarize, in terms of fidelity to English linguistic lexical patterns, these random lexicon groups are expected to follow the following hierarchy:

$$\text{uniform} < \text{Zipfian} < \text{scrambled} < \text{bigram} < \text{trigram}$$

Comparison of these random lexicons with each other and with the real English lexicon allows us to determine which observed properties of English are lexically meaningful. If a property is true of all PNNs, it is likely to be a simple consequence of the definition of the neighbourhood relation over lexicons, and does not necessarily reveal anything about language. If a property is true of the real English PNN and the n-gram PNNs, but not the other random PNNs, it is likely to be a consequence of the phonotactic patterns of the lexicon—hard limits on what shapes words can take. If a property is true only of the English PNN but not any of the random PNNs, then it is likely to be due to a deeper organizing principle of the lexicon.

2.2 Network measures

For each group of PNNs, several network measures were taken.

- Giant component size: the size, as a ratio of the number of vertices in the entire graph, of the largest connected component.
- Clustering coefficient: the mean clustering coefficient for each vertex in the entire graph.
- Mean number of neighbours: the mean number of neighbours for each vertex in the entire graph.

[4] Even in borrowed words like *Ndebele*, a short vowel sound is usually inserted before the /n/.

- Assortativity by degree [13]: the correlation coefficient of the degree of a vertex with that of its neighbours, averaged over the entire graph. This measures the extent to which highly-connected words cluster together.
- Shortest path: the average shortest path length for all pairwise comparisons. Vertices which are not connected are ignored, essentially yielding a grand mean of each connected component weighted by the number of vertices in each component.

2.3 Robustness to vertex removal

To evaluate the relative robustness of each PNN, vertex removal was performed. A proportion of vertices were removed, and the average shortest path of the graph was measured. The procedure was then repeated with a larger proportion of vertices. This procedure allows us to examine the change in the robustness of the network as successively more vertices were removed.

Two vertex removal methods were employed: a random method, where vertices were removed at random; and a targeted method, where vertices were removed in decreasing order of degree. That is, the word with the most neighbours was removed first, the word with the second most was removed second, and so on. We tested removal proportions from 0 to 0.05, in 21 equally-spaced steps. Two measures of network robustness were used: giant component size and average shortest path. We follow convention in assuming that larger giant component size and smaller shortest path represent more robust networks.

3 Results

Table 1 summarizes the results for the real English PNN and of the five groups of random PNNs.

3.1 Overall patterns

For giant component size, clustering coefficient, and mean number of neighbours, the statistics obeyed the following hierarchy:

$$\text{uniform} < \text{Zipfian} < \text{scrambled} < \text{bigram} \approx \text{trigram} \approx \text{English}$$

That is, the n-gram PNNs were very similar to the real English PNN, while the other random PNNs had lower values as a function of their projected similarity to English. Nevertheless, while the other random PNNs were not similar to English, their statistics do indicate some small-world properties, as previously reported [1, 20].

The size of the real English PNN giant component is still smaller than most scale-free networks studied in the literature [14]. The fact that the real English PNN regardless has the largest giant component of all the PNNs suggests that the

English lexicon has clusters of highly-connected words [18]. For this to happen, the lexicon must employ a large degree of *re-use* of common elements and sequences of phonemes. It has been theorized that such re-use is beneficial for the developing lexicon in infant and child language acquisition [3], and aids in the processes of speech production and perception in adults [4, 5].

All the PNNs examined are assortative by degree: words with many neighbours tend to cluster together. Assortativity was higher for the *n*-gram random PNNs than the other random PNNs, and it was highest of all for the English PNN. Taken together, these results suggest that the property of assortativity in general is intrinsic to PNNs, but that it is enhanced by the presence of phonotactics, and enhanced further by unknown lexical organizational constraints.

The real English PNN had neither the longest nor the shortest mean shortest path length. This value does not appear to readily distinguish the real English PNN from the random PNNs, nor does it distinguish the different random PNNs from each other.

Table 1: Summary statistics for the real English PNN and the five groups of random PNNs. Standard deviations included in parentheses. GC: giant component; Clust.: clustering; Sh.: shortest.

	GC size	Clust. coefficient	Mean # neighbours	Assortativity	Sh. path
Uniform	.023 (.002)	.009 (.001)	0.108 (0.010)	.540 (.045)	6.032 (0.334)
Zipfian	.100 (.003)	.034 (.002)	0.628 (0.032)	.240 (.021)	4.835 (0.073)
Scrambled	.167 (.002)	.046 (.001)	0.710 (0.010)	.427 (.019)	7.057 (0.091)
Bigram	.286 (.004)	.106 (.002)	2.604 (0.050)	.459 (.009)	5.242 (0.038)
Trigram	.371 (.005)	.138 (.002)	3.018 (0.055)	.538 (.008)	6.432 (0.068)
English	.320	.117	2.675	.643	6.991

3.2 Vertex removal

The patterns of robustness to vertex removal are shown in Fig. 3 for giant component size, and Fig. 4 for average shortest path length. For all groups of PNNs, random vertex removal does not appear to influence giant component size, while targeted vertex removal leads to a decline in giant component size. However, it can be seen that the fall is very sharp for the uniform, Zipfian, and scrambled PNNs (rapidly reaching zero), while the slope is much gentler for the bigram, trigram, and real English PNNs.

The same pattern is observed for the shortest path length: no change for random removal, rapid increase for targeted removal for the uniform, Zipfian, and scrambled PNNs, and gentle increase for targeted removal for the bigram, trigram, and real English PNNs. After a point, the shortest path lengths for the uniform, Zipfian, and scrambled PNNs fall; this is a consequence of the rapid fragmenting of the graph into many isolated islands, and does not reflect an increase in robustness. (Note that the falls coincide with the giant component size approaching zero.)

These results demonstrate that, while the real English PNN is remarkably robust to both random and targeted vertex removal [1], the same is true of the bigram and trigram random PNNs. The observed robustness is therefore not necessarily due to an organizing principle of lexical structure, but phonotactic limitations on possible words.

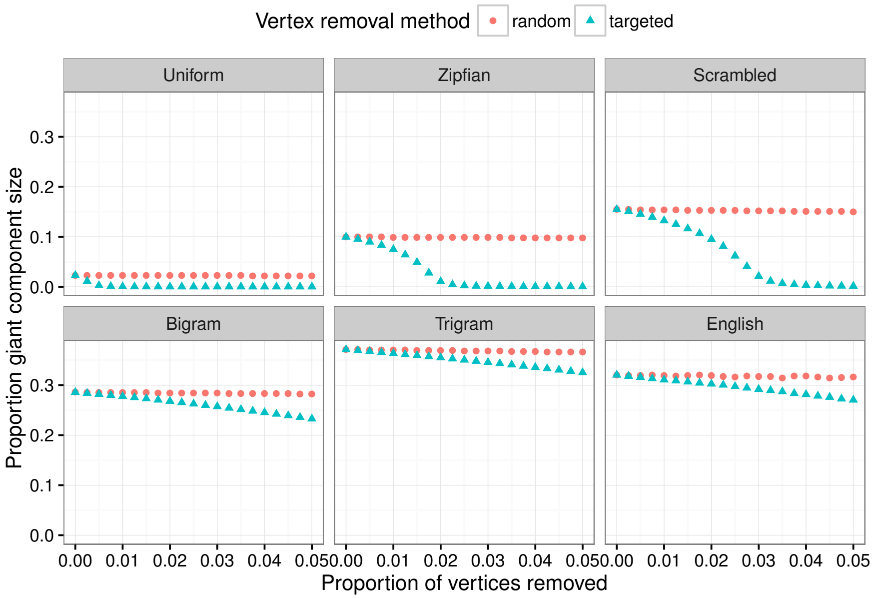

Fig. 3: Giant component size for the five random groups of PNNs, plus the real English PNN, given two vertex removal methods, plotted as a function of the proportion of vertices removed. Red circles depict values for random vertex removal; blue triangles depict values for targeted vertex removal.

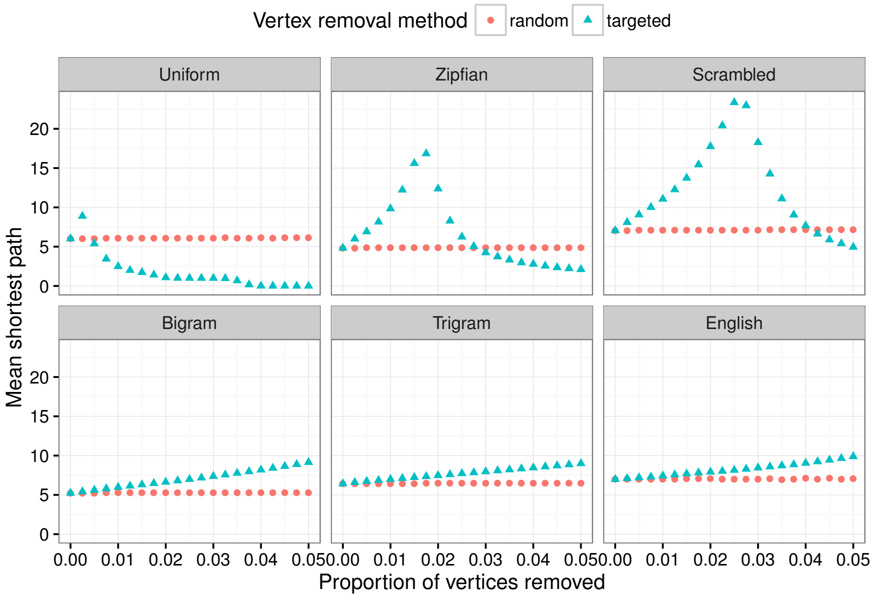

Fig. 4: Shortest path lengths for the five random groups of PNNs, plus the real English PNN, given two vertex removal methods, plotted as a function of the proportion of vertices removed. Red circles depict values for random vertex removal; blue triangles depict values for targeted vertex removal.

4 Discussion

For both the real English PNN and the random PNNs, the clustering coefficients were relatively high, confirming the assertion that PNNs have small-world properties [1]. However, as this was observed for the random PNNs too, it would appear to be a property intrinsic to the definition of a PNN, and therefore not necessarily psycholinguistically meaningful.

In terms of giant component size and mean number of neighbours, the real English PNN was midway between the bigram and trigram random PNNs, suggesting that these properties are due to phonotactics rather than any deeper constraints which may modulate the development of the lexicon.

However, where the real English PNN stood out from the random PNNs was in assortativity by degree. While all the PNNs were assortative, the real English PNN was the most of all. It is possible that this high level of assortativity aids in lexical retrieval by limiting the spread of activation to irrelevant candidate words in the process of speech perception [20]. However, the mechanisms by which the real English PNN obtains this high level of assortativity is unknown.

Finally, the vertex removal analysis demonstrated that while the real English PNN and the *n*-gram PNNs were very robust to targeted vertex removal, the other random PNNs rapidly lost robustness. In this regard, the non-*n*-gram random PNNs are similar to scale-free networks, in that the mean shortest path length rapidly increases upon targeted vertex removal [14]. This finding suggests that the robustness observed by [1] is not necessarily due to a particular cognitive constraint on lexical organization, but a consequence of phonotactics.

5 Conclusion

With a novel method for generation of random PNNs, we have shown that some properties of PNNs—such as small world properties, small giant component size, and assortativity by degree—are due to the definition of the neighbourhood relation that defines PNNs, rather than properties of language *per se*. Others properties are common to the real PNN and n-gram PNNs, which simulate the phonotactic patterns of natural language. For example, the n-gram PNNs are indistinguishable from the real PNN in terms of giant component size, clustering coefficients, and mean number of neighbours, and all are equally robust to vertex removal. These properties are likely due to phonotactics, rather than the definition of the neighbourhood relation or any underlying cognitive constraints.

A promising avenue for further study is the strong assortativity observed on the real PNN relative to the random PNNs, suggesting that there could be principles and mechanisms governing the structure of the lexicons of human languages which enhance the assortativity of the network. Whether these principles operate over milliseconds (i.e. they are caused by patterns of cognitive processing) or generations (i.e. they are caused by patterns of cultural evolution) is a promising question for future research. Replicating these results for languages other than English is also a crucial step in establishing the true nature of PNNs.

Acknowledgements This work has received support under the program "Investissements d'Avenir" launched by the French Government and implemented by ANR with the references ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

References

[1] Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010). The structure of phonological networks across multiple languages. *International Journal of Bifurcation and Chaos*, 20(03), 679–685.

[2] Blake, I., & Gilchrist, J. (1973). Addresses for graphs. *IEEE Transactions on Information Theory*, 19(5), 683–688.

[3] Carlson, M. T., Bane, M., & Sonderegger, M. (2011). Global properties of the phonological networks in child and child-directed speech. In *Proceedings of the 35th Boston University Conference on Language Development* (Vol. 1, pp. 97–109). Somerville, MA: Cascadilla Press.

[4] Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 35(6), 1934–1949.

[5] Chan, K. Y., & Vitevitch, M. S. (2010). Network structure influences speech production. *Cognitive Science*, 34(4), 685–697.
[6] Deza, M., & Grishukhin, P. (1993). Hypermetric graphs. *The Quarterly Journal of Mathematics* Oxford (2), 44, 399–433.
[7] Deza, M., & Laurent, M. (1994) ℓ_1-rigid graphs. *Journal of Algebraic Combinatorics*, 3, 153–175.
[8] Deza, M., & Shpectorov, S. (1996). Recognition of the ℓ_1-graphs with complexity $O(nm)$, or football in a hypercube. *European Journal of Combinatorics*, 17(2), 279–289.
[9] Graham, R. L., & Winkler, P. M. (1985). On isometric embeddings of graphs. *Transactions of the American Mathematical Society*, 288(2), 527–536.
[10] Gruenenfelder, T. M., & Pisoni, D. B. (2009). The lexical restructuring hypothesis and graph theoretic analyses of networks based on random lexicons. *Journal of Speech, Language, and Hearing Research*, 52(3), 596–609.
[11] Imrich, W., & Klavžar, S. (1997). Recognizing Hamming graphs in linear time and space. *Information Processing Letters*, 63(2), 91–95.
[12] Mahowald, K., Dautriche, I., Gibson, E., Christophe, A., Piantadosi, S. T. (In revision). Lexical clustering in efficient language design.
[13] Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E*, 67(2), 026126.
[14] Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2), 167-256.
[15] Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, Indiana University*, 10, 357–376.
[16] Shoemark, P., Goldwater, S., Kirby, J., & Sarkar, R. (2016). Towards robust cross-linguistic comparisons of phonological networks. In *Proceedings of the 14th ACL SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, 110.
[17] Shpectorov, S. V. (1993). On scale embeddings of graphs into hypercubes. *European Journal of Combinatorics*, 14(2), 117-130.
[18] Siew, C. S. (2013). Community structure in the phonological network. *Frontiers in Psychology*, 4, 553.
[19] Stella, M., & Brede, M. (2015). Patterns in the English language: phonological networks, percolation and assembly models. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(5), P05006.
[20] 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.
[21] Zörnig, P., & Altmann, G. (1983). The repeat rate of phoneme frequencies and the Zipf-Mandelbrot law. *Glottometrika*, 5, 205–211.