PHONOLOGICAL NETWORK PROPERTIES OF NONWORDS INFLUENCE THEIR LEARNABILITY

Rory Turnbull

Newcastle University
rory.turnbull@newcastle.ac.uk

ABSTRACT

It is well-established that a word's phonological neighbourhood is relevant for many aspects of phonetic processing. The current study investigated whether this is true of nonwords too. In the exposure phase, participants (*N*=198, English-speaking) were shown pairings of novel words and novel items. In the test phase, participants were presented with a word–item pairing and asked if the item was named correctly or not. Overall accuracy was high (79.2%). Nonwords were learned with higher accuracy, and responded to with faster reaction times, if they had more neighbours, and if their neighbours were more well-connected, relative to nonwords which had fewer neighbours and whose neighbours were less well-connected. These results demonstrate that existing words in the lexicon can exert an influence on the acquisition and processing of novel words.

Keywords: lexical effects, network structure, word learning, acquisition

1. INTRODUCTION

Several studies have shown that the phonological neighbours of a word—all the words that differ from that word only in terms of a single phoneme—can influence that word's processing, both in terms of perception and production [1]. This finding has been extended, to some extent, to the processing of nonwords too. For example, lexical decision reaction time task is modulated by the number of neighbours of nonword stimuli [2, 3], and nonword neighbourhoods can influence phoneme identification [4].

It is less clear, however, how neighbourhood properties affect the learnability of novel words. By definition, a word which has not yet been learned is functionally equivalent to a nonword. This paper examines how neighbourhood properties influence novel word learning.

Figure 1: Example novel item stimulus image.

2. METHOD

2.1. Materials

Wuggy [5] was used to generate several hundred British English nonwords, of which 47 were selected for use in the experiment. The stimuli were recorded by a native speaker of British English. Images of 48 novel items with low familiarity ratings were selected from the NOUN database [6]. An example image is provided in Figure 1. The nonword stimuli were as follows:

bɒsp	bjuːf	bjuːp	bjuːʃ	
bjuːʒ	bjʊə	bræb	bræmpt	bræv
ðiːd	ðiːpt	druːʤ	eɪf	fɜːʧ
flɔːʧ	fruːm	fwɑːθ	gjuːʃ	glæg
glaɪb	gleɪm	grʌlpt	hjuːf	huːls
jɜːft	klɒm	krɪnθ	kwɛmpt	pɜːb
plɛsp	pwiːθ	skwaɪs	skwɒf	slæʤ
sləʊg	smɪd	spliːʃ	splʌnd	stʌg
swɒsk	ʃnæm	ʃpiːg	traʊd	vjuːt
zɛft	θrɒmps	θwæʃt		

2.2. Neighbourhood properties

Several recent studies conceive of the phonological lexicon as a complex network, where words, represented as nodes in a network, are connected

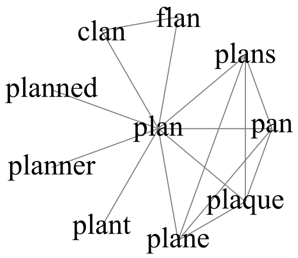

Figure 2: Example phonological neighbourhood network centred around the English word *plan*. Note that some neighbours of a word are neighbours of each other. Adapted from [10].

to their phonological neighbours [7, 8]. This yields a web, part of which is visualised in Figure 2. This approach allows for the examination of more nuanced measures than just the raw number of neighbours.

Four neighbourhood properties are considered in this study [9]. The **degree** of a node is the number of neighbours it is connected to. This is also referred to as "neighbourhood density". A node's **coreness** is a measure of how well-connected its neighbours are. The coreness of a node is the largest value of k for which the node is in the k-core of the network; the k-core of a network is the subset of network with only the nodes with degree of k or more. The **clustering coefficient** of a node is the proportion of the neighbours of a node which are neighbours of each other. **Closeness centrality** is a measure of how close a node is to others in the network, defined as the reciprocal of the sum of the distance from that node to all other nodes in the network.

Some stimulus words are **singletons**. These are nodes in the network with a degree of zero, that is, no neighbours. Singletons and non-singletons are analysed separately due to the fact that the neighbourhood properties for singletons are all either zero or undefined.

Each of these measures was calculated for each stimulus word, and also for the nonword's lexical neighbours, using the CELEX lexicon of UK English. As each stimulus varies in how many neighbours it has, the values for the neighbours were averaged. For example, the nonword /bɒsp/ has a degree of 4, and its neighbours have an average degree of 13.25. The python package Networkx was used to calculate these values [11].

These measures are partially correlated with each other. Some of this correlation is due to the definitions of the measures themselves—for example, coreness is intrinsically linked to degree—but some are due to properties of phonological lexicons in general. For example, there tends to be a correlation between the degree of a word and the degree of its neighbours—this 'assortativity by degree' has been noted to be a property of lexical organisation in natural language [12, 13]. Table 1 shows a correlation matrix of the network measures of the non-singleton stimuli.

Due to this collinearity, these measures are inappropriate to be used simultaneously in a regression analysis. A factor analysis was undertaken to reduce the dimensionality of the neighbourhood measures. Horn's parallel analysis returned two factors as optimal; factor loadings are reported in Table 2. Factor 1 corresponds largely to the coreness and degree of the nonword stimulus and the mean degree of its neighbours. Factor 2 corresponds largely to the closeness centrality of both the nonword stimulus and its neighbours.

2.3. Procedure

The experiment consisted of 6 blocks, each with an exposure phase and a test phase. In the exposure phase, a novel item was presented on screen and the item's name (a nonword stimulus) was presented auditorially. Eight item–word pairs were presented;

		Word			Neighbour			
		Coreness	Degree	Closeness	Clustering	Coreness	Degree	Closeness
Word	Clustering	0.571	0.459	0.384	-0.113	0.390	0.379	0.342
	Coreness		0.955	0.572	-0.119	0.686	0.740	0.435
	Degree			0.560	-0.144	0.631	0.681	0.398
	Closeness				0.068	0.713	0.692	0.922
Neighbour	Clustering					-0.084	-0.169	0.196
	Coreness						0.960	0.705
	Degree							0.659

Table 1: Correlation matrix of the network measures for the non-singleton stimuli and their neighbours.

		Factor 1	Factor 2
Word	Clustering	0.566	0.108
	Coreness	0.997	0.008
	Degree	0.957	-0.013
	Closeness	0.569	0.753
Neighbour	Clustering	-0.126	0.272
	Coreness	0.685	0.457
	Degree	0.739	0.381
	Closeness	0.429	0.900

Table 2: Factor loadings for the two factors returned from the factor analysis of the non-singleton neighbourhood measures.

each pair was presented four times, in random order. In the test phase, participants were again presented with an item–word pair and asked if the word was the item's correct name. Participants responded via keyboard, with no time limit on their responses. The experiment was constructed in Labvanced and participants were recruited via Prolific, and took roughly 10 minutes to complete.

2.4. Predictions

As singletons have no neighbours, they are less similar to existing words in the lexicon than non-singletons. Singletons should therefore be harder to integrate into the lexicon than non-singletons, and as such we should observe lower accuracy on these words.

For the other neighbourhood properties, extremely high or extremely low values are expected to inhibit learning. For example, if a new word has an extremely high clustering coefficient, it is being integrated into a very 'busy' neighbourhood. A very low clustering coefficient is a very 'sparse' neighbourhood and should be similarly challenging. Only words in the 'Goldilocks zone' should be learned easily. This prediction is necessarily vague as we do not currently have the empirical data to know what values of the neighbourhood properties count as "extremely high" or "extremely low".

2.5. Participants

The study was completed by 198 participants, all of whom were UK residents fluent in English. Most of the participants (191) were native speakers of English.[1] Participants were compensated GBP2.25 for their participation.

2.6. Analysis

Four regression models were constructed to model the data: a logistic mixed effects regression model to predict response accuracy in singleton versus non-singleton stimuli; a linear mixed effects regression model to predict reaction time (RT) in singleton versus non-singleton stimuli; a generalised additive mixed model (GAMM) [14] to predict accuracy from the neighbourhood factors; and a GAMM to predict RT using the neighbourhood factors. The RT models included data only from correctly-answered trials in which the picture and word matched.

The singleton versus non-singleton models included fixed effects of the stimuli's status as a singleton or not, the overall position of the trial in the experiment (trial number), and their interaction. Random intercepts of participant and word were included. The neighbourhood factors are not predicted to have necessarily linear effects, which motivated the use of GAMMs. Factors 1 and 2 were included as thin plate regression smooths. A parametric term of trial number was included. Random intercepts of participant and word were also included.

3. RESULTS

3.1. Singleton versus non-singleton nonwords

A speed–accuracy tradeoff was observed in the models comparing singleton to non-singleton words. As participants proceeded through the experiment, they responded overall faster ($t = -4.637$, $p < .001$) and less accurately ($z = -3.136$, $p = 0.002$). No other significant effects were observed: singleton words did not appear to be significantly harder or easier to learn than non-singleton words.

3.2. Neighbourhood properties of non-singleton words

The output of the GAMM predicting response accuracy is summarised in Table 3. Mirroring the results of the earlier analysis, accuracy decreased as trial number increased. A significant effect of Factor 1 was observed, such that stimuli with higher Factor 1 scores (higher coreness, higher degree, and higher-degree neighbours) were responded to more accurately than stimuli with lower Factor 1 scores. This effect is visualised in Figure 3.

The output of the GAMM predicting RT is summarised in Table 4. RT decreased with trial number, reflecting the speed–accuracy tradeoff mentioned in the previous section. Significant

Parametric coefficients:			
	Estimate	*t*	*p*
Intercept	2.483	17.934	< .001
Trial number	-0.019	-4.258	< .001
Approximate significance of smooth terms:			
	EDF	*F*	*p*
Factor 1	1	4.225	0.040
Factor 2	1	3.634	0.057

Table 3: Model output from the GAMM predicting response accuracy.

Parametric coefficients:			
	Estimate	*t*	*p*
Intercept	7.591	513.596	< .001
Trial number	-0.001	-4.128	< .001
Approximate significance of smooth terms:			
	EDF	*F*	*p*
Factor 1	1	13.514	< .001
Factor 2	1	4.244	0.039

Table 4: Model output from the GAMM predicting log reaction time.

effects of both Factor 1 and Factor 2 were observed, visualised in Figure 4. Words with higher Factor 1 scores (higher coreness, higher degree, and higher-degree neighbours) and higher Factor 2 scores (higher closeness centrality) were responded to faster than words with lower Factor 1 and 2 scores.

4. DISCUSSION

The results do not support the hypothesis that non-singleton words are learned better than singleton words: both singleton and non-singleton words had similar response accuracies and RTs. However, there were differences among the non-singleton words, governed by neighbourhood properties. Notably, words with higher degree or higher coreness, and which had higher-degree neighbours, were responded to faster and more accurately than words with lower values. Similarly, words with high closeness centrality (and high-centrality neighbours) were responded to more quickly than words with lower values.

These findings are consistent with a network-based model of the lexicon, where activation spreads from neighbour to neighbour [15, 16, 8]. The high-centrality words can be accessed faster than the low-centrality words, as they are closer to the core of the lexicon and therefore more accessible. That neighbourhood properties affect (non)word learning is a novel finding. This study represents a step towards a better understanding of the kinds of lexical influences on word learning, and may provide insights on the lexical organisation of languages, first language acquisition, and second language acquisition.

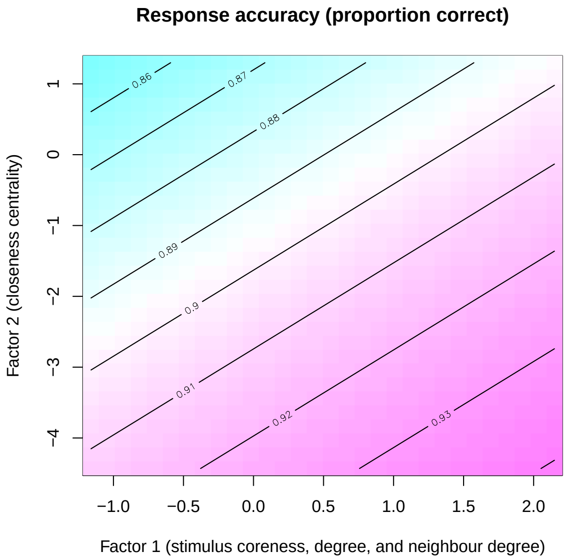

Figure 3: Effects of Factor 1 and Factor 2 on response accuracy.

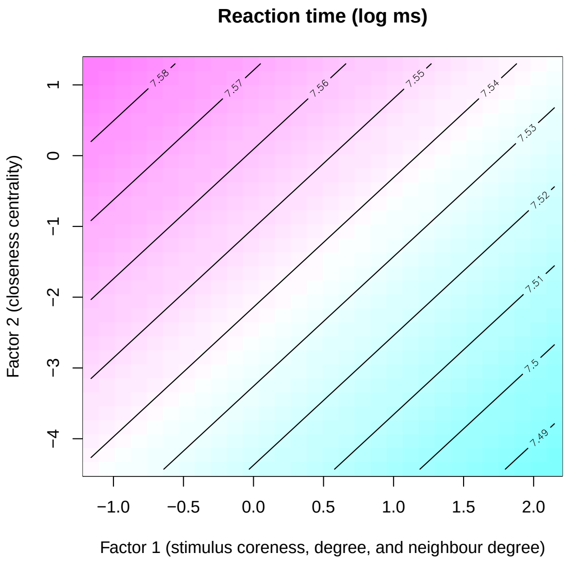

Figure 4: Effects of Factor 1 and Factor 2 on reaction time.

ACKNOWLEDGEMENTS

I am grateful for comments and suggestions from the Newcastle University Phonetics & Phonology Research Group and two anonymous reviewers.

5. REFERENCES

[1] M. S. Vitevitch and P. A. Luce, "Phonological neighborhood effects in spoken word perception and production," *Annual Review of Linguistics*, vol. 2, no. 1, pp. 75–94, 2016.

[2] E. Janse, "Neighbourhood density effects in auditory non-word processing in aphasic listeners," *Clinical linguistics & phonetics*, vol. 23, no. 3, pp. 196–207, 2009.

[3] M. C. Kelley and B. V. Tucker, "The recognition of spoken pseudowords," *Language, Cognition and Neuroscience*, pp. 1–22, 2022.

[4] R. S. Newman, J. R. Sawusch, and P. A. Luce, "Lexical neighborhood effects in phonetic processing." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 23, no. 3, p. 873, 1997.

[5] E. Keuleers and M. Brysbaert, "Wuggy: A multilingual pseudoword generator," *Behavior Research Methods*, vol. 42, no. 2, pp. 627–633, 2010.

[6] J. S. Horst and M. C. Hout, "The novel object and unusual name (noun) database: A collection of novel images for use in experimental research," *Behavior research methods*, vol. 48, no. 4, pp. 1393–1409, 2016.

[7] M. Stella and M. Brede, "Patterns in the English language: Phonological networks, percolation and assembly models," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 5, p. P05006, 2015.

[8] M. S. Vitevitch, "What can graph theory tell us about word learning and lexical retrieval?" *Journal of Speech, Language, and Hearing Research*, vol. 51, pp. 408–422, 2008.

[9] M. Newman, *Networks*, 2nd ed. Oxford, UK: Oxford University Press, 2018.

[10] R. Turnbull, "Graph-theoretic properties of the class of phonological neighbourhood networks," in *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*. Online: Association for Computational Linguistics, Jun. 2021, pp. 233–240. [Online]. Available: https://aclanthology.org/2021.cmcl-1.27

[11] A. Hagberg, P. Swart, and D. S Chult, "Exploring network structure, dynamics, and function using NetworkX," in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Pasadena, CA, 2008, pp. 11–15.

[12] S. Arbesman, S. H. Strogatz, and M. S. Vitevitch, "The structure of phonological networks across multiple languages," *International Journal of Bifurcation and Chaos*, vol. 20, no. 3, pp. 679–685, 2010.

[13] R. Turnbull and S. Peperkamp, "What governs a language's lexicon? determining the organizing principles of phonological neighbourhood networks," in *Complex Networks & Their Applications V*, ser. Studies in Computational Intelligence, H. Cherifi, S. Gaito, W. Quattrociocchi, and A. Sala, Eds. Cham, Switzerland: Springer, 2017, vol. 693, pp. 83–94.

[14] S. N. Wood, *Generalized Additive Models: An Introduction with R*, 2nd ed. Boca Raton, FL: Chapman & Hall / CRC, 2017.

[15] M. Goldrick and S. E. Blumstein, "Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters," *Language and Cognitive Processes*, vol. 21, no. 6, pp. 649–683, 2006.

[16] M. Goldrick, H. R. Baker, A. Murphy, and M. Baese-Berk, "Interaction and representational integration: Evidence from speech errors," *Cognition*, vol. 121, pp. 58–72, 2011.

[1] Removal of the non-native speakers from analysis did not affect the interpretation of the results.