Phonological Networks and Systematicity in Early Lexical Acquisition

Catherine E. Laing[1]

[1] University of York, York, UK

Author Note

All code and associated data for this manuscript can be found on the project's OSF page at https://osf.io/uzrsy/?view_only=340858d2084245d087fc00fcca41b679. This study was not pre-registered.

Correspondence concerning this article should be addressed to Catherine E. Laing, Department of Language and Linguistic Science, University of York, Heslington, YO10 5DD. E-mail: catherine.laing@york.ac.uk

Abstract

Infants' early words tend to be phonologically similar. This may reflect a systematic approach to early production, as they adapt newly-acquired forms to fit familiar structures in the output. This 'rich-get-richer' approach to phonological acquisition, known as *preferential attachment* in network science, proposes that new words cluster together with existing phonologically-similar words in the lexicon (or network). This contrasts with recent work (e.g. Fourtassi et al., 2020) showing that the learning environment is the key predictor in learning (*preferential acquisition*). This study expands on previous analyses of vocabulary norm data to analyse naturalistic data, namely phonetic transcriptions of nine infants' word productions, from word onset to age 2;6. Network growth models test whether 1) acquisition is best modeled through preferential attachment or preferential acquisition, 2) the trajectory of network growth changes over time, and 3) there are any differences in network growth of adult target forms vs. infants' actual productions. Results show that preferential attachment predicts acquisition of new words more convincingly than preferential acquisition: newly-acquired words are phonologically similar to existing words in the network. Furthermore, systematicity becomes increasingly apparent over the course of acquisition, and infants produce their early words more systematically than we would expect from looking at target forms alone.

Keywords: systematicity, phonological development, preferential attachment, networks analysis

Phonological Networks and Systematicity in Early Lexical Acquisition

Table 1

Outputs from nested model comparisons comparing logistic regression models predicting acquisition of words in each month according to INT- and EXT-like growth structures.

	Actual			Target		
Model	Df	Chi Sq	p	Df	Chi Sq	p
null vs. INT	2	395.48	<0.001	2	84.18	<0.001
null vs. EXT	2	1.00	0.608	2	2.87	0.238
INT vs. INT+EXT	2	0.35	0.841	2	0.33	0.848
EXT vs. INT+EXT	2	394.83	<0.001	2	81.64	<0.001

Table 2

Results from maximal logistic regression model (model 3) testing the effects of network growth values, corpus (English as baseline), input frequency, comprehensive AoA, word category and word length to predict word acquisition. All variables were scaled and centred. Category has been removed for ease of interpretation but this is shown in the full model output in S4.

	Actual					Target				
Effect	beta	SE	z	p	95% CI	beta	SE	z	p	95% CI
Intercept	-3.28	0.27	-12.25	<0.001	[-3.81,-2.76]	-2.96	0.38	-7.79	<0.001	[-3.71,-2.22]
INT value	0.88	0.07	12.59	<0.001	[0.74,1.02]	0.47	0.06	8.08	<0.001	[0.35,0.58]
EXT value	0.01	0.05	0.27	0.787	[-0.09,0.11]	0.00	0.05	0.04	0.964	[-0.1,0.1]
Age	0.94	0.14	6.66	<0.001	[0.66,1.21]	1.25	0.13	9.77	<0.001	[1,1.5]
AoA	-0.23	0.04	-6.22	<0.001	[-0.3,-0.15]	-0.21	0.04	-5.79	<0.001	[-0.29,-0.14]
Length	-0.08	0.06	-1.46	0.144	[-0.2,0.03]	-0.14	0.06	-2.32	0.021	[-0.25,-0.02]
Input freq	0.17	0.05	3.30	0.001	[0.07,0.27]	0.19	0.05	3.71	<0.001	[0.09,0.29]
Corpus	0.43	0.31	1.37	0.172	[-0.18,1.03]	0.93	0.47	1.99	0.047	[0.01,1.84]
Age x INT	0.16	0.05	2.96	0.003	[0.05,0.27]	-0.07	0.04	-1.62	0.106	[-0.16,0.01]
Age x EXT	-0.03	0.04	-0.58	0.561	[-0.11,0.06]	-0.03	0.05	-0.56	0.575	[-0.12,0.06]
Age x AoA	0.12	0.03	3.97	<0.001	[0.06,0.17]	0.10	0.03	3.28	0.001	[0.04,0.16]
Age x Length	0.03	0.05	0.72	0.474	[-0.06,0.13]	0.01	0.05	0.22	0.825	[-0.09,0.11]
Age x Input freq	-0.06	0.03	-1.90	0.057	[-0.12,0]	-0.06	0.03	-1.92	0.055	[-0.13,0]