Phonological Networks and Systematicity in Early Lexical Acquisition

Catherine E. Laing[1]

[1] University of York, York, UK

Author Note

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

Decades of work on phonological development has documented the systematic nature of infants' earliest words. Studies of phonetic (McCune & Vihman, 2001) and phonological structures (Vihman, 2016) show that many of a child's first word forms share similar properties. Infants draw on what they know: when articulatory, memory and planning capacities are simultaneously limited, a "phonic core of remembered lexical items and articulations" (Ferguson & Farwell, 1975, p. 112) may help them deal with the challenge of developing an early lexicon. Vihman (2019, p. 263) describes the early lexicon as "an emergent network of related forms" that develops systematically, in line with the well-rehearsed segments and structures already in the infant's inventory. A networks approach to phonological development offers one way of identifying and quantifying this systematicity. In this study, I present a longitudinal analysis of nine infants' lexical development to identify systematicity in the first three years of word production. I consider the phonological characteristics of the developing lexicon using network analysis to demonstrate how early systematicity may support infants to acquire the requisite capacity for flexible and automatic word production.

In early development, the combined challenges of articulation, memory and planning mean that the constraints on infants' production are high, and so they draw on a limited set of vocal outputs that represent a growing number of target words. According to Vihman (2014, 2019), word production begins with a small lexicon of phonologically-simple and accurately-produced forms, which are 'selected' for their ease of production, as well as their perceptual salience. As the lexicon grows, target forms that do not necessarily fit these structures are 'adapted' so that they do. Selection of and adaption to accessible phonological structures indicate the presence of systematicity within the developing lexicon. Essentially, the new target form is allocated to one of a small number of accessible or well-rehearsed motoric categories, and as these categories increase in size they become

increasingly entrenched (Thelen & Smith, 1996). In data from their bilingual (English-Spanish) daughter's early word acquisition, Deuchar and Quay (2000) show that 13 of her first 20 words are produced with a CV structure, and many are phonologically identical: she produces *car*, *clock*, *casa* 'house' and *cat* as [ka], and *papa* 'daddy', *pájaro* 'bird' and *panda* as [pa]. This demonstrates a 'pattern force', whereby production is driven by a small number of well-rehearsed structures. This tendency to acquire similar-sounding forms may continue throughout development: Mitchell, Tsui and Byers-Heinlein (2022) show that French-English bilingual infants are more likely to acquire translation equivalents that are similar in phonological form (cognates, e.g. *banana* and *banane*) than non-cognate word pairs (e.g. *dog* and *chien*) upto age 27 months. Systematicity in phonological acquisition may thus support lexical development over the first three years.

One way of interrogating systematicity in early phonological productions is through network analysis, which offers a quantitative perspective on the organization and development of the lexicon. Developmental research in this area centres around the words that children target in production to establish connectivity on phonological (e.g. Siew & Vitevitch, 2020) and semantic (e.g. Hills, Maouene, Maouene, Sheya, & Smith, 2009) planes. That is, how similar target words are to one another in form or meaning, and what this might mean for acquisition. However, as yet there is no work looking at the way children *produce* those words; that is, whether or not children are drawing on systematicity in the output. Given the extensive background research that suggests a systematic approach to early word *production*, expanding network analysis to this area is a natural next step for language development networks.

The term *network* refers to a web of forms (or *nodes*, in network terms) that are interconnected based on shared properties. Here these are phonological properties, but could also be semantic, or indeed non-linguistic properties such as genetic information, social connections or location (see Bell et al., 2017, for a review). Network growth models

analyse changes within a system over time, and two key models[1] of development have been proposed for lexical acquisition: preferential attachment (hereafter **INT** due to the assumption that network growth is internally-driven in this model; note that some studies refer to this model as PATT) and preferential acquisition (hereafter **EXT**, due to the assumption that network growth is externally-driven in this model, note that this model is otherwise known as PAQ, Hills et al., 2009; see also Steyvers & Tenenbaum, 2005). INT models of network growth propose a rich-get-richer scenario, whereby the most highly-connected nodes (nodes with more *edges*) in the network are most likely to attract new nodes. In phonological development terms, this model implies that the lexicon will constitute clusters of similar-sounding words (i.e. denser phonological neighbourhoods), and that a child is more likely to acquire new words that attach to these dense clusters: infants' production of newly-acquired words will be similar to their production of existing words in the lexicon. INT-like growth is therefore driven by the *internal* linguistic system. On the other hand, EXT-like growth assumes that forms that connect to (i.e. share properties with) a higher number of *different* nodes in the target network will be acquired first. EXT models of network growth thus assume that *external* factors in the learning environment influence acquisition – that is, forms that are most well-connected within the target language will be acquired earlier. In phonological terms, this would mean that early productions would constitute the distribution of segments and structures that co-occur most frequently in the input, thus leading early forms to resemble the statistical properties of the ambient language more closely, rather than a 'pattern force' driven by dominant features of the existing lexicon. For example, given an existing lexicon that included the forms *pat* and *bat*, an INT model would predict that a highly phonologically-similar form

––––––

[1] A third model - Lure of the Associates - predicts that new words will be learned that are similar to the highest number of already-known words in the network. This model has been considered in some studies (Hills et al., 2009; Siew & Vitevitch, 2020) but will not be considered here as there is no conclusive evidence for this model in the development literature, though note that there is evidence for this model in adult word learning (e.g. Stamer & Vitevitch, 2012; Storkel, Armbrüster, & Hogan, 2006).

such as *pit* or *bit* might be acquired next, whereas EXT would predict that more variable forms would be acquired, such as /p/-initial or /t/-final words, which have high phonotactic probability in English and thus connect to a wider range of different forms.

Existing studies show mixed evidence for INT- and EXT-like growth[2] in lexical development. Hills and colleagues' (2009) study of semantic networks showed evidence for EXT, but not INT, in associative networks of normed vocabulary acquisition data. Amatuni and Bergelson (2017) support this with an analysis of a large-scale corpus of input data combined with normed productive vocabulary data derived from WordBank (Frank, Braginsky, Yurovsky, & Marchman, 2017). These same approaches have also been applied to phonological data: Fourtassi, Bian and Frank (2020) analyse both phonological and semantic network growth from vocabulary norms (receptive and productive) in 10 languages to find consistent evidence in support of EXT-like growth, for both phonological and semantic networks, receptive and productive vocabularies, and across the 10 languages included in their analysis. Ciaglia, Stella and Kennington (2023) analysed complex multiplex networks (including phonological, semantic, sensorimotor and visual associations) to find evidence for both EXT and INT in word learning, though evidence was stronger for EXT. In contrast, Siew and Vitevitch (2020) tested phonological networks in acquisition of older Dutch- and English-learning children (age 3-9 years), again using vocabulary norms to indicate age of acquisition for each word. Their analysis revealed contrasting findings for English compared with Dutch, as well as an age effect: INT-like network growth predicted acquisition in English and Dutch, and both EXT and a third model (Lure of the Associates, see note above) predicted word learning in Dutch. INT was a better predictor of acquisition earlier on in development (i.e. earlier-acquired words were likely to attach to densely-connected clusters of similar forms); later on, the opposite was found, whereby later-acquired words tended to be phonologically more distinct (i.e. less similar to existing words in the network). A replication of this study using data from adult second-language

———

[2] Note that these are not mutually exclusive.

learners of English found consistent results (Luef, 2022). Evidence in favour of INT has also been found in adult word-learning experiments: for example, Mak and Twitchell's (2020) work with paired-association learning in adults shows that participants were better at remembering word pairs when items had been paired with cue words in semantic space that had a higher degree (i.e. were connected to a larger number of semantically-similar words). The authors propose that these highly-connected words may support learning due to the fact that they tend to be used more flexibly, and thus occur in a more diverse set of linguistic contexts. In infancy, this relates back to Ferguson and Farwell's "phonic core of remembered lexical items and articulations" (1975, p. 112), as infants apply the same well-rehearsed phonological form flexibly and systematically to new items in the lexicon.

These studies present an intersection of evidence for the role of INT and EXT network growth in phonological development. However, two key aspects of these existing approaches should be expanded further. First, the consideration of acquisition in terms of only target forms provides no view of systematicity in *production*, which is where systematicity has been most well-documented in naturalistic data. Second, vocabulary norming data abstracts away from the individual differences expected in early phonological development (e.g. Vihman, Kay, Boysson-Bardies, Durand, & Sundberg, 1994); by drawing on data that generalises across hundreds (or even thousands) of children, it may not be possible to capture developing systematicity due to individual differences in the words and sounds that are acquired first. This makes it difficult to test which model of network growth (INT or EXT) is most cogent. To better understand the role of systematicity in early word production, it is essential to consider infants' *actual productions* of their early word forms, in terms of both *how* and *when* they produce them. In this paper, I analyse phonological networks of both *target* and *actual forms* (that is, the words children produce, and the way they produce them) produced in naturalistic data from two languages, in order to consider phonological systematicity within the individual development trajectories of nine infants.

Hypotheses

Drawing on naturalistic data, this study uses network growth models to capture phonological connectivity (taken here as an index of systematicity) within the individual lexicons of nine infants. Two sets of networks will be established for each infant: one tracing connections between infants' *actual* word productions, the other between the *target* productions of these forms. Network analysis will quantify systematicity in the developing lexicon via two key network growth frameworks: INT and EXT. I will draw on approaches outlined in previous studies (Amatuni & Bergelson, 2017; Fourtassi et al., 2020; Siew & Vitevitch, 2020) to test whether naturalistic data reveals evidence of systematicity in infants' output forms, such that language development is shaped by existing production knowledge. Specifically, I predict that:

H1) Developing phonological networks will show stronger evidence of an INT-like model of growth over an EXT-like model, based on evidence from the phonological development literature that shows phonological similarity across individual infants' lexicons (e.g. Vihman & Keren-Portnoy, 2013).

H2) INT-like growth will be most evident earlier on in development, as infants select and then adapt words to fit their production capacity (Vihman, 2019). Later, more variability is expected as phonological capacity develops. This would also align with previous evidence for INT-like growth in toddlers (Siew & Vitevitch, 2020) and novel word learning in adults (Mak & Twitchell, 2020).

H3) If INT-like growth is supported in the data, then this should be more convincing for Actual than Target productions, given that we expect infants to adapt target words to fit the motor routines that are most accessible to them in production. This difference is not expected for an EXT-like model of network growth, which assumes that network growth reflects connectivity in the input; the question of differences between Actual and Target forms is thus not of central theoretical interest for EXT models in this analysis.

To test these hypotheses, phonological networks will be established for nine infants acquiring American English or French. Phonological distance will be calculated between each word and each other word in each infant's network to establish connectivity within the network. Logistic regression models and Generalized Additive Mixed Models (GAMMs) will determine whether acquisition of *Actual* and *Target* forms reflects INT- or EXT-like growth in early phonological development, and how these networks change over time.

Methods

This analysis follows approaches taken by Hills and colleagues (2009) and Siew and Vitevitch (2020), by establishing network growth values for each word in each child's lexicon. Logistic regression models will be used to test whether INT or EXT growth values can best predict word learning. This is followed with the use of Generalized Additive Mixed Models (GAMMs) to analyse the trajectory of network growth values over time.

Data and Materials

Data for this study was extracted from CHILDES (Child Language Data Exchange System, MacWhinney, 2000) using Phon (Hedlund & Rose, 2020). Two corpora were selected for the analysis: American English (Providence corpus, Demuth, Culbertson, & Alter, 2006) and French (Lyon corpus, Demuth & Tremblay, 2008). These corpora were selected due to their parallel data collection and transcription methods. The English data includes five infants (including two boys[3]) and four from the French corpus (two boys).

---

[3] The Providence corpus (Demuth et al., 2006) includes six children (three boys). One child was later diagnosed with a developmental disorder and so is omitted from this analysis. The Lyon corpus (Demuth & Tremblay, 2008) includes five children (two boys) but one of the datasets (Marilyn) is not fully transcribed and is therefore excluded from this analysis.

Both corpora include spontaneous[4] interactions between child and caregiver, recorded in the home for one hour every two weeks from the onset of first words. The original corpora were orthographically transcribed, and then phonetically transcribed and checked by trained coders. See Demuth et al. (2006) and Demuth and Tremblay (2008) for full details of data collection and annotation.

Transcripts were extracted from the first session in the dataset (the first session in which the child produced a word) until age 2;6. Data was analysed on a month-by-month basis, such that all new word types produced in each month were aggregated to give a rolling monthly network of all words produced by each child. The session in which a word first occurred was considered the session in which it was 'acquired', and was included in that month's list of newly-acquired words. Later productions of the same word were not included in the dataset. Two of the American infants (Naima and Lily) had denser data taken at weekly intervals during some periods of data collection, but this is not considered to be an issue as no between-child comparisons will be conducted, and subject will be coded as a random effect in all statistical models. The total network of words at any given month amounts to all the unique words produced up to and including that month. All tokens of each newly-acquired word produced by each infant in each session were extracted (*Actual* forms, i.e., the phonological form as produced by the child) alongside their target transcription (*Target* forms).

Only words included on the US English and French communicative development inventories (CDIs, Fenson et al., 1994; Kern & Gayraud, 2010) were analysed. Following Jones and Brandt (2019), every unique word was considered, though plurals were categorised with their singular nouns. For example, *fall*, *fell* and *falling* were considered as unique words (coded under the CDI 'basic level' *fall*), while *bananas* was categorised with

---

[4] Note that while interactions were naturalistic and thus not at all directed by the original researchers, the data was not coded for infant productions that were imitated from or prompted by the caregiver, and so data includes both spontaneous and non-spontaneous infant productions.

its singular form *banana* and *children* with *child*. In the French data, this rule was also applied to masculine/feminine forms: *animaux* was categorised with the singular *animal*, and feminine *petite* was categorised with masculine *petit*. Words with the same basic level form that were orthographically different but phonologically indistinguishable (e.g. verb forms in French, such as *aime* and *aiment* from the infinitive *aimer* 'to love') were categorised together. This approach was taken in order to account for developmental changes in infants' word production (i.e. the production of more complex morphological forms) while also avoiding coding two words as different that share almost identical forms and meanings (e.g. plural nouns). To generate networks of Actual and Target forms, phonological distance was calculated between every word and every other word in the cumulative network at each month, following Monaghan, Christiansen, Farmer and Fitneva's (2010) approach. This is based on phonological features, following Harm and Seidenberg (1999) and based on Chomsky and Halle's (1968) theory of government phonology. This was considered to be the most appropriate measure of phonological distance, as oppose to other established measures such as Levenshtein distance in phonemes (e.g. Fourtassi et al., 2020; Siew & Vitevitch, 2020): distinctive features allow us to consider distance on a phonologically-appropriate gradient, whereby the difference between words such as *bat* and *pat* is smaller than the difference between *bat* and *rat*. Using edit distance (in phonemes) as a measure, *pat*, *bat* and *rat* would be equidistant, thereby equating all phonemes as articulatorily similar, which does not reflect the reality of phonological development: /p/ and /b/ are among the earliest consonants to be acquired, whereas /r/ is not typically acquired until around age 5 (cf. McLeod & Crowe, 2018). Note that in the present analysis only consonants were included, given that vowels are highly variable in production until around age 3, and notoriously difficult to transcribe from child speech (Donegan, 2013; Kent & Rountrey, 2020). This means that two words that differ only in their vowel segments are coded as the same in the current analysis. Words were aligned by syllable nucleus: onset consonants were compared with other onset consonants,

and codas were compared with codas. Full criteria for establishing distance, alongside tabulated examples, are included in the Supplementary Information, S1. When multiple tokens of the same word type were produced in a single session, the values derived from the distinctive feature matrix were averaged across tokens to create a mean phonological representation for each word type. While this is not a perfect measure, it captures a metric of both variability and similarity within and between each word type.

Altogether, 5483 words were excluded from the data due to not appearing on the French or American English CDIs (2224 in French and 3259 in English). The final dataset includes 3096 word types overall, aggregated across infants (English=1933, French=1163). On average, there were 32 tokens of each word type (SD = 144); 3 words occurred only once in the data, and on average each word type was produced across 6 different months (SD = 8), which supports the (admittedly imperfect) assumption made here that the first production of a word in the dataset indicates its acquisition in the infant's lexicon. See Table 1 for a breakdown of the data by corpus and child. All but 10 tokens (all French) in the data had three syllables or fewer in the target form, with 1 syllable on average in the English data (SD = 0.50) and 1.53 in the French data (SD = 0.66).

Network Analysis

For each child, two kinds of network were generated: 1) a *global network*, which represents the final network, i.e. all words produced in the data by 2;6. This network includes the Target production of all individual word types produced in the dataset, coded for age of first production. The global network is taken to reflect the learning environment, or the input, which is why only Target forms are included; this will be used to establish EXT growth values for each word in the data (see below), and also serves as a proxy for the 'end-state' towards which each child's phonological development is directed. 2) A series of 'known' networks representing the lexicon at each month. Each monthly known network includes all the words produced up to and including the given month, in either Actual (the

infants' realization) or Target (the target realization) form. This series of networks is used to generate INT values for each word in the data. As a reminder, for both kinds of networks, a given word type was included from the first session in which it occurred, and multiple tokens of a given word type in that session were 'averaged out' to one unique set of distinctive feature values for each word, from which connectivity with all other word types was then derived. Connectivity was established between all words in the global network, and all words in each monthly network; two nodes were considered to be connected (i.e., formed an edge) if they had a scaled phonological distance of 0.25 or less; this value was established by standardizing all phonological distance values for each infant, and then selecting the the value that captured the first quartile of connectivity within the dataset. The first quartile of connectivity across individual corpora (English and French) and data types (Actual and Target) ranged from a scaled distance of0.18 to 0.24 (see S2 and S3 for further details); the threshold of 0.25 thus represents the upper limit of the total variability across the four subsets of data. All edges in the networks were unweighted and undirected.

Once networks were established, INT and EXT values were calculated for each word. Following Siew and Vitevitch's (2020) approach, these values were generated by computing, for each month, the likelihood that an as-yet-unknown word (i.e. all the words in the global network - that is, all words produced in a given child's data up to and including age 2;6 - that had not yet been produced) would form an edge with known words in the existing network (i.e. the words produced up to and including a given month). The INT value of a given yet-to-be-learned word represents the median degree of all the words it would connect to (i.e. those with a phonological distance of 0.25 or less) if it were learned in the following month. For example, a word with an INT value of 5.6 would connect to a set of words in the following month that, on average, connected to 5.6 other words each. Given that INT assumes that newly-acquired words will connect to already-well-connected words in the existing network, higher INT values predict learning in the following month: new words will connect to words with higher median degrees. INT networks were generated with both

Actual and Target forms. EXT values reflect the degree of a given word in the global network of all words produced by 2;6. So a word with an EXT value of 87 connects to 87 other words in the global network. See Figure 1 for a visualization of these two models of acquisition. Again, as EXT predicts that well-connected words in the global network would be acquired earlier, higher EXT values predict earlier learning; in each month, we would expect that as-yet-unknown words with the highest EXT values will be acquired in the following month. As EXT-like growth is assumed to represent the connectivity of words in the ambient language, global networks were established with Target forms only, since the way infants produce words in the existing network is not relevant to this model. However, given that connectivity differs across Target and Actual networks (i.e., the known words in the Actual network at month n may be different from the known words in the Target network in the same month), both Actual and Target network structure will be tested in the analysis. To clarify, as both INT and EXT values are established through connectivity in the network (i.e. only words that form an edge with another word are represented), the words included in each network differs slightly; 54 words did not connect to any other word at a threshold of 0.25 in the Actual data, and 63 words in the Target data. For the same reason, the size of Actual (n = 3266) and Target (n = 3257) networks differs, as some forms connected at a threshold of 0.25 in their Actual, but not their Target, forms.

Data Analysis

Network growth models.   Network growth models will be used to address the first two hypotheses. Network growth models are logistic regression models that predict whether or not a word is learned in the following month; the dependent variable is whether or not a word was learned in month $n+1$ (learned vs. not learned). The key predictors of acquisition are INT/EXT growth values for each word at each month. The models test the assumption that higher growth values predict earlier learning, such that words with higher INT/EXT values at month n are more likely to be learned at month $n+1$. Following

predictions set out in H1, model comparisons should show INT values to be a better predictor of word learning than EXT values. H2 predicts age-related changes in the effect of INT; an INT x Age interaction is expected to show INT to be a better predictor of learning at earlier time-points.

GAMMs.   It is also a possibility that any age-related changes will be non-linear. To address this, Generalized Additive Mixed Models (GAMMs) will be used to test H2, following Wieling (2018) and Sóskuthy (2017). GAMMs allow analysis of dynamically varying data (i.e. change over time), without assuming change to be linear. Since there is no clear expectation as to whether any age-related changes would be linear or not, testing H2 using both logistic regression and GAMMs will account for both possibilities. Non-linearity in the data is analysed in the model through the inclusion of *smooth terms* and *random smooths*, which capture the non-linearity of fixed and random effects, respectively, alongside parametric terms. The dependent variable in these models will be INT and EXT values (tested as predictors in the network growth models outlined above); if predictions set out in H2 are borne out in the data, then we would expect to see a significant effect for age on INT/EXT values as a smooth in the model. H3 will also be tested using GAMMs, given that any differences between Actual and Target data may change over time. Here, we would expect to see a significant effect for Data type as a parametric term. These effects will be identified through nested model comparison and inspection of smooth plots. Full model details are provided below.

All code for data preparation and analysis can be found on the project's OSF page at https://osf.io/uzrsy/ (Laing, 2024). This study was not pre-registered.

Results

Age of production (AoP) ~ connectivity.   First, to assess the broader assumption that connectivity in the network will change systematically over time, regardless of whether that is through INT- or EXT-like changes, the relationship between

age of production (AoP) and connectivity (degree) in the static network was considered. Both INT and EXT models of network growth assume that later-acquired words will be less well-connected in the network. Across all infants, there was a mean AoP~degree correlation of r=-0.21 (Spearman's, SD=0.09; English: r=-0.26, SD =0.04; French: r=-0.15, SD=0.11); overall, later-learned words were less well-connected in the networks. Negative correlations were found in all children's data except Anais, and these were all significant at p<.05 except Anais and Nathan (French corpus). See S3. This is consistent with previous similar work showing that earlier-learned words are more highly-connected in the network (Fourtassi et al., 2020; Hills et al., 2009), and replicates these findings with a naturalistic sample of infant production data. To ascertain that this negative relationship between AoP and connectivity is not simply a given in vocabulary-based networks that increase in size over time, this analysis was re-run on an identical dataset that was randomized by AoP, such that new words were added to the French and English networks at random ages, and then the degree of each word in this random network was calculated. Across the data, there was no correlation between AoP and degree (r=0.01; p=.678); evidence for INT/EXT-like growth in the real data is thus not an inevitable outcome of vocabulary growth.

Network growth models. Next, network growth models were generated to test whether INT and EXT values predicted which words were produced in the following month. As a reminder, models of both INT- and EXT-like acquisition predict that, for each month, the as-yet-unknown words with the highest INT/EXT values should be learned in the following month. Logistic mixed effects regression models included a binomial dependent variable (coded as 0 or 1) indexing whether, for each as-yet-unknown word at month *n*, it was acquired in month *n+1*. As well as INT and EXT growth values, each model also included target word length in phonemes, reported age of acquisition for each item in the comprehensive vocabulary according to vocabulary checklists (CDIs; see below), input frequency in child-directed speech, word category (based on CDI word categories), and corpus (English vs. French) as fixed effects. Infant was specified as a random effect

with a by-infant random slope for the effect of age. Input frequency for each word was derived from Braginsky et al.'s (2019) frequency estimates, which includes a unigram count for every word produced in adult speech in all CHILDES corpora for the respective language. Normed comprehensive vocabulary data for English and French CDI words was taken from WordBank (Frank et al., 2017); again following Braginsky and colleagues (2019), age of acquisition (AoA) was taken as the month in which >50% of children were reported to understand a given word. As comprehensive vocabulary norms are only available up to ages 16/18 months for French/US English data, respectively, in total 1470 tokens did not include this measure (603 word types across all infants), either due to the word being acquired after the cut-off age for the CDI checklist (i.e. it was included on the checklist but fewer than 50% of infants understood the word by 16/18 months), or due to it not being included on the checklist in the first place (i.e. it is included on the productive vocabulary checklist only). All relevant variables were scaled and centered; INT/EXT values were scaled by speaker and age to account for the effect that increased vocabulary size at each month has on INT/EXT values (i.e. when the network is bigger, a newly-acquired word has the opportunity to connect to a higher number of different words by default). In each case interactions were included between Word length x Age, Word frequency x Age, AoA x Age and INT/EXT values x Age. P-values were established through nested model comparisons. Analysis of Actual/Target data includes INT values for the Actual/Target network, respectively. EXT values always represent connectivity in the Target network (i.e. to simulate the adult production of a given word in the input), but models were run on both Actual and Target data since connectivity in these data sets differed due to words generally being more highly connected at a threshold of 0.25 in the Actual data than the Target data, as explained above. These models were run using the *lme4()* package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2020).

Following Siew and Vitevitch (2020), the first step was to construct three models: a null model (model 0) with word length, input frequency, comprehensive AoA, word

category, corpus, and age included as predictors of word learning, and then two additional models with INT (model 1) and EXT (model 2) growth values included as additional predictors, respectively. Models 1 and 2 were then compared against model 0 to test for the effects of INT and EXT values individually. A third model (model 3) was then constructed that included both INT and EXT values as predictors. Data type (Actual and Target) was modeled separately in each case. The full model specification for model 3 is as follows:

Model 3: Learned next ~ EXT value * Age + INT value * Age + Word length * Age + Input frequency * Age + AoA (comprehension) * Age + Category + Corpus + (1 + Age|Speaker)

In the Actual and Target data, INT values improved model fit over and above the effects of input frequency, comprehensive age of acquisition, word length, category, corpus, and age, whereas EXT values did not. See Table 2. When EXT values were added to the model testing just INT values, model fit was not improved over and above the effects of INT alone, but when INT values were added to the model testing only EXT values, model fit improved in both Actual and Target data. INT thus appears to be a predictor of acquisition in both Actual and Target data, while EXT does not appear to predict learning.

Model outputs are shown in Table 3. In both Actual and Target data, higher INT values predicted acquisition (Actual data: b=0.88, $p < .001$; Target data: b=0.47, $p <$ .001), providing support for H1. Alongside INT values, age, input frequency and comprehensive AoA were all significant predictors of acquisition in both Actual and Target data: unsurprisingly, words were more likely to be acquired at higher ages, and words with a lower comprehensive AoA (according to vocabulary norms for each language) were more likely to be learned, as were words that were more frequent in caregiver speech. Word length and corpus were also significant in the Target data only: words were more likely to be acquired in the following month in the English data, likely because the English corpus was larger than the French corpus (see Table 1), and as we might expect, shorter words

were more likely to be learned. Category has been removed the Table 3 for ease of reading, but is shown in the full model output in the SI (S4). AoA interacted significantly with age in both Actual and Target data: as one might expect, words with lower comprehensive AoA norms were typically learned at earlier timepoints, and vice versa. The INT x Age interaction was significant in the Actual data only; contrary to predictions, INT values increased over time. This was not dependent on increasing vocabulary size as increasing network size was controlled for in the scaled INT variable.

INT-like growth over time. H2 predicted a change in INT-like growth over time, such that INT values should predict learning more effectively in earlier acquisition than later acquisition. That is, earlier words should have higher INT values relative to vocabulary size than later-acquired words. As reported above, a significant INT x Age interaction was observed in only the Actual data, and the effect of age on phonological systematicity is not as expected: INT values of newly-learned words are lower earlier on in development (see Table 3).

To explore these results further, GAMMs were run using the *mgcv()* package in R (Wood, 2011), to observe how INT values changed over time as new words were learned. The data was subsetted to include only INT values at the time-point immediately prior to the word's production as the dependent variable in the model (i.e. for a word produced at 17 months, its INT value at 16 months was analysed); higher INT values are expected to predict that a word would be learned in the next month. This left 2766 data points for the Actual data, and 2702 for the Target data. EXT values for the same month were included as a fixed effect. This time, INT/EXT values were scaled only by speaker, not by age, in order to more clearly visualize the data, though note that results were consistent when the values from the previous models were included (see S6). Otherwise, models incorporated the same fixed effects and interactions as in the mixed-effects regression models above. By-infant and by-corpus random smooths were included in the model for the effect of age; these control for by-infant and by-corpus differences in the data over time. To account for

the fact that adjacent values (i.e. INT values at month *n* and month *n+1*) were likely correlated, an autocorrelation parameter was included, which was derived from an initial full model. The start point for each infant's dataset (i.e. their first recording session) was also indexed in the model. To test for the effect of age, model comparisons were run using the *compareML()* function from the *itsadug()* package (Rij, Wieling, Baayen, & Rijn, 2022): the full model included the effect of age as a smooth term, as well as interactions between age and EXT values, input frequency, word length, and AoA. This was compared to another model that did not include the effect of age in either smooth terms or interactions. Because model summaries for GAMM smooths may be non-conservative (Sóskuthy, 2017), any significant effects in the initial model comparisons will be assessed using smooth plots of the models. To continue to explore the independent roles of INT- and EXT-like growth in the data, the same models will also be run with EXT values as the dependent variable (and INT values as a fixed effect). This component of the analysis will be exploratory given that we have no expectation as to how EXT values will affect learning over time, and given that EXT values did not significantly predict learning. As above, Actual and Target data were modeled separately.

Outputs from model comparisons are shown in Table 4 (rows 1-2). Consistent with findings from the logistic regression models above, age had a significant effect on INT values in the Actual data, and not in the Target data. While results from model comparisons are to be treated with caution for GAMMs, model smooths for INT show a convincing linear change in INT values with age. See Figure 2. Consistent with the regression model coefficients above, and again contrary to the expectations set out in H2, in both the Actual and Target data, INT values were lower in earlier acquisition, and increased over time. Model smooths for the EXT values are shown in Figure 3 for comparison purposes, where we see a moderate decrease in EXT values over time.

Data type comparisons. H3 predicted that systematicity would be stronger in Actual, compared to Target, data. We would therefore expect INT values to be higher in

Actual data overall, indicating more connectivity. This analysis only applies to INT, given that the global network used to determine EXT-like growth is generated from Target forms anyway; the expected substantial overlap in the two data types is shown in Figure 3. To test for an effect of Data type, GAMMs were used to account for any non-linearity in the data over time. Model structure was almost identical to that reported above, except that 1) Data type was included as a parametric term, with a difference smooth[5] and a by-Data type random smooth for the effect of age; 2) the full dataset, incorporating Actual and Target forms together, was tested.

Results from a nested model comparison are shown in Table 4 (row 3). Data type had a significant effect on INT values. A summary of the full model reveals that INT values were significantly lower in the Target data than the Actual data (b=-0.66, $p<$ .001), thereby supporting H3.

The difference of the two smooths is shown in Figure 4. The red line indicates periods where the two trajectories differed significantly from one another - from ~15 months until the final time-point in the analysis. For clarity, the two smooths are visualised in Figure 5 where the difference between the two trajectories is apparent. A visualization of how the data differs across infants is shown in S6.

Discussion

This study tested two established frameworks of network growth in the context of early phonological development: preferential attachment (INT) and preferential acquisition (EXT) (Fourtassi et al., 2020; Hills et al., 2009; Siew & Vitevitch, 2020). Using naturalistic data to observe infants' realization of words, it was possible to establish similarity (or

---

[5] Difference smooths account for the fact that the different levels of the smooth might differ in their non-linearity; in this instance, the by-Data type difference smooth accounts for the possibility that Actual and Target data may have different trajectories.

connectedness) across the phonological properties of infants' early words, and map how this changes over time. Based on previous analyses showing that infants' early productions tend to share phonological properties (e.g. Vihman, 2016; Waterson, 1971; see also Vihman & Keren-Portnoy, 2013), it was hypothesised that the early vocabulary would grow in an INT-like manner (H1) – that is, it should constitute dense clusters of similar-sounding forms – and that acquisition should be most systematic earlier on in development (H2). Expanding on two key studies in this area (Fourtassi et al., 2020; Siew & Vitevitch, 2020), it was also predicted that a network consisting of infants' actual productions (that is, the child's *realization* of the target forms) should demonstrate more typical INT-like growth than an equivalent network constituting just the target forms (H3). Two of these three hypotheses were supported by the data.

First, in support of H1, network growth models showed strong evidence for INT-like growth in both Actual and Target data; newly-acquired words were produced in a similar way to existing words in the network, such that, in a given month, as-yet-unknown words that would connect to the most densely-clustered known words were more likely to be acquired in the next month. EXT-like growth did not predict learning in any of the models. H2 predicted that INT-like network growth would be stronger in earlier development, based on previous analyses that show infants' earliest words to be phonologically similar or even identical (e.g. Deuchar & Quay, 2000). However, the opposite was true in this data set: in both Actual and Target data, earlier-acquired words tended to have *lower* INT values, while later-acquired words had higher INT values. Finally, in support of H3, INT-like growth was more convincing for the Actual than the Target data: analysis of GAMM smooths revealed that Data type (Actual versus Target) accounted for significant variance in INT values, whereby Target data had significantly lower INT values than Actual data from early on in the data (15 months).

It was surprising to find no evidence for EXT across the analyses, given that previous studies show more convincing evidence for EXT overall, and given that INT and EXT are

not mutually exclusive models of network growth. Amatuni and Bergelson (2017) propose that INT and EXT could work together, such that EXT may "[supplement] INT by providing a structured sampling space for new word selection" (p.5). That is, a combination of INT and EXT would provide both internal (output-driven) and external (input-driven) roles in development. Indeed, acquisition is a dynamic and interactive process (Thelen & Smith, 1996), with ample evidence showing the effects of the input on early word learning (Ambridge, Kidd, Rowland, & Theakston, 2015; Rowe, 2012); it is to be expected that both models would be at work simultaneously during acquisition. It may be that this was not shown in the current data due to the fact that the regression models controlled for many external factors known to affect word learning – input frequency, word length, word category, etc. – which together could have accounted for much of the variability that otherwise would have been captured by EXT growth values in this corpus. It may also be the case that the representation of the Target network was not sufficiently aligned with the reality of the end-state network that the infants will acquire. Analysing the Target network on a larger scale – for example, including all words produced in the infants' inputs across their recordings, and building a network based on which of these words infants produce in the dataset – might better represent the role of EXT-like growth on early word learning. This is an avenue that could plausibly be considered in future work.

The present analysis sheds new light on systematicity in early language acquisition, specifically regarding the role of INT- and EXT-like models of phonological development. Previous studies have drawn on age of acquisition data, using the target form as the index of production (Fourtassi et al., 2020; Siew & Vitevitch, 2020). This has allowed study of vocabulary growth across a large sample, and findings have presented a new perspective on the role of phonological neighbourhoods in early acquisition. However, these analyses have not interrogated the role of *production*. By considering networks in relation to the way infants produce their early-acquired words, it has been possible to consider phonological network growth from a novel perspective. The findings presented here reveal a systematic

approach to early phonological development, as infants exploit their existing production capacity to produce new words with familiar articulatory routines. These results support many previous studies that show lexical development to take place via the implementation of systematic structures and templates (Vihman, 2019; Vihman & Keren-Portnoy, 2013; Waterson, 1971), and also model a new way of analysing phonological systematicity in infants' early productions, which can be extended to larger samples and applied to a wider variety of languages.

Given that Fourtassi and colleagues (2020) analysed data from children of similar ages using the same subset of words (i.e. CDI words), we would expect the current findings to map on to their results, particularly in the analysis of Target data. However, their study consistently reveals stronger evidence for EXT and so our results do not align. This may reflect direct differences in the type of data used: in the present study, the order of acquisition (and thereby the model of network growth) reflects the chronological order of individual children's production. Month-by-month acquisition norms taken from thousands of children's CDIs model an 'average' order of acquisition, whereby words that *tend to* appear earlier in the developing lexicon are biased towards an earlier age of acquisition. Frank and colleagues (2021) report the first 10 words of infants acquiring American English, which (for stop consonants only) contain two instances of /m/, three each of /n/ and /d/, five /b/ and one /g/. In naturalistic production, however, a word's phonological form may prime the acquisition of other similar-sounding words: production of *baby* may be shortly followed by *bib* and *ball* (cf. McCune & Vihman, 2001), while in vocabulary norms, acquisition of *baby*, *bib* and *ball* is represented at the group level. Vocabulary norming data thus represents an 'averaging out' of phonological connectedness across thousands of infants, creating a bias towards EXT-like growth. Previous similar studies perhaps represent a more general, one-size-fits-all trajectory to lexical development, whereas these results capture individual clusters of connectivity as children acquire words that match the phonological characteristics of existing words in the lexicon.

Indeed, studies of infants' early words show that, on a word-by-word basis, early-acquired forms tend to consist of the same set of consonants, in both target and actual forms. This reflects the child's 'selection' of early words to match their own consonant repertoire (McCune & Vihman, 2001; Stoel-Gammon & Cooper, 1984; Vihman, 2019). Given that these results show evidence for INT-like growth in both Actual and Target data, it appears that infants are selectively acquiring forms that match their own production preferences, and are either producing these forms accurately (*selected*, in Vihman's terms) or *adapting* them to match their preferred output patterns. Within Vihman's framework, phonological development involves the selection or adaption of lexical units to fit a set of easily-accessible articulatory categories. That is, an infant systematically acts upon new understanding (i.e. acquired receptive vocabulary items) within the limitations of their development, selecting existing categories to deal with challenges presented in production. These are 'well-worn paths' that represent the stable and well-rehearsed production routines that drive selection, and later adaption, of infants' early word forms. In producing forms that are accessible and familiar to the child, they can 'rehearse' particular segments and structures, easing up memory and planning capacity for more flexible and variable production further down the line.

These results align with and expand on previous work observing phonological neighbourhood density (PND) and phonotactic probability in early word learning. Both have been found to positively influence new word acquisition earlier on in development (Coady & Aslin, 2003; Dollaghan, 1994; Storkel, 2004), though for older children (Charles-Luce & Luce, 1990) and adults (Gordon & Kurczek, 2014; Vitevitch & Luce, 1999), low neighbourhood density appears to be more beneficial in learning and remembering novel words. The present findings suggest that, at least in early development, high PND (i.e. phonologically more similar words in the lexicon) may in part be derived from systematicity in production. That is, if infants are selecting new words that match their output capacity in early development, then we would expect a higher number of

phonological neighbours in the Target and Actual forms, as observed here, and as consistent with the PND literature. On the other hand, the fact that INT predicts acquisition in both Target and Actual forms may be due to the increased learnability of words that belong to denser neighbourhoods, leading infants to produce these earlier on – the fact that they are also phonotactically similar (due to PND and phonotactic probability being correlated, Vitevitch & Luce, 1999) would no doubt support their early production as infants need to draw on fewer resources to produce a number of new words. Results in this study lend preference to the first explanation (i.e. that higher PND is motivated by production, rather than the other way around): we see a continuous increase in Actual INT values over time as new words are adapted to fit existing well-rehearsed segments and structures (i.e. existing dense neighbourhoods attract phonologically similar words for acquisition), which is significantly higher than Target INT values over the same period (see Figure 5). If higher PND was motivated by learning, we would expect to see no difference in acquisition of Actual and Target forms, since infants would be learning words that clustered together just as densely in the Target network as in the Actual network, i.e. they wouldn't be systematically adapting words to fit the dominant patterns and structures in their existing lexicon.

This study raises new questions for future analyses into systematicity in phonological development. While efforts were made to fully characterise the phonological content of infants' early productions – through using distinctive features with Euclidean, rather than Levenshtein, distance, and observing Actual productions alongside Target forms – still it was not possible to capture the full extent of systematicity, i.e. the presence of prosodic structures or templates (Vihman, 2019). Future work in this area should expand the analyses to consider the development and systematic implementation of templates. Furthermore, this analysis considers only two languages; it would be valuable to extend the approach to a wider variety of languages. Systematicity has been demonstrated across languages (Arnon & Clark, 2011; Khattab & Al-Tamimi, 2013; Szreder, 2013), and so it

should be possible to find cross-linguistic commonalities in network growth. Typological differences in network growth would raise questions about the cognitive reality of systematicity in phonological development. The influence of semantic networks on acquisition has also not been considered here - further studies may want to analyse similar naturalistic data to consider semantic network growth within infants' actual productions, or even combine indices of semantic connectivity with that of phonological networks to observe how/whether the two interact in early development. Finally, it would have been valuable to have data on these infants' comprehensive vocabularies over the course of the analysis. While comprehensive vocabulary norm data was included in the models, this is a wide step away from the expectation posited throughout this paper that individual trajectories shape learning. Comprehensive vocabulary data would allow an analysis of the extent to which known (but not yet produced) words "fit" existing segments and structures in the child's productive repertoire; in this way, models could be devised that predict which words in the comprehensive vocabulary are most likely to appear next in the productive vocabulary.

Conclusion

When naturalistic data is considered within a phonological networks account, we find evidence for INT-like network growth, but not EXT-like growth. English- and French-learning infants acquired words that would connect to the most highly-connected nodes in the existing network (INT-like growth), and this became increasingly systematic over time. When we look at the target form of the words infants acquire and how they produce them, in both cases we see evidence to show that early acquisition is driven – at least in part – by preferences in the output. That is, infants acquire words that cluster together phonologically, and produce them systematically such that early production represents clusters of similar-sounding forms.

References

Amatuni, A., & Bergelson, E. (2017). *Semantic networks generated from early linguistic input* [Preprint]. Animal Behavior; Cognition. https://doi.org/10.1101/157701

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*(2), 239–273. https://doi.org/10.1017/S030500091400049X

Arnon, I., & Clark, E. V. (2011). *Experience, variation and generalization: Learning a first language.* John Benjamins Publishing.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bell, M., Perera, S., Piraveenan, M., Bliemer, M., Latty, T., & Reid, C. (2017). Network growth models: A behavioural basis for attachment proportional to fitness. *Scientific Reports*, *7*(1), 42431. https://doi.org/10.1038/srep42431

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and Variability in Children's Word Learning Across Languages. *Open Mind*, *3*, 52–67.

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, *17*(1), 205–215. https://doi.org/10.1017/S0305000900013180

Chomsky, Noam, & Halle, Morris. (1968). *The sound pattern of english.* New York, NY: Harper; Rowe.

Ciaglia, F., Stella, M., & Kennington, C. (2023). Investigating preferential acquisition and attachment in early word learning through cognitive, visual and latent multiplex lexical networks. *Physica A: Statistical Mechanics and Its Applications*, *612*, 128468. https://doi.org/10.1016/j.physa.2023.128468

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, *30*(2), 441–469. https://doi.org/10.1017/S0305000903005579

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of english. *Language and Speech*, *49*(2), 137–174.

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of french determiners. *Journal of Child Language*, *35*(1). Retrieved from https://www-cambridge-org.abc.cardiff.ac.uk/core/journals/journal-of-child-language/article/prosodicallyconditioned-variability-in-childrens-production-of-french-determiners/B34D994A1EA63DFF841D1C2C7974D6FE

Deuchar, M., & Quay, S. (2000). *Bilingual acquisition: Theoretical implications of a case study.* Oxford, UK: Oxford University Press.

Dollaghan, C. A. (1994). Children's phonological neighbourhoods: Half empty or half full? *Journal of Child Language*, *21*(2), 257–271. https://doi.org/10.1017/S0305000900009260

Donegan, P. (2013). Normal vowel development. In M. J. Ball & F. E. Gibbons (Eds.), *Handbook of vowels and vowel disorders* (pp. 24–60). Psychology Press.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, M., Stephen J. Tomasello, . . . Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*. https://doi.org/10.2307/1166093

Ferguson, Charles. A., & Farwell, C. B. (1975). Words & sounds in early language acquisition. *Language*, *51*, 419–439.

Fourtassi, A., Bian, Y., & Frank, M. C. (2020). The growth of children's semantic and phonological networks: Insight from 10 languages. *Cognitive Science*, *44*(7),

e12847. https://doi.org/10.1111/cogs.12847

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language, 44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The wordbank project.* MIT Press.

Gordon, J. K., & Kurczek, J. C. (2014). The ageing neighbourhood: Phonological density in naming. *Language, Cognition and Neuroscience, 29*(3), 326–344. https://doi.org/10.1080/01690965.2013.837495

Hedlund, G., & Rose, Y. (2020). *Phon 3.1 [computer software].* Retrieved from https://phon.ca

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks preferential attachment or preferential acquisition? *Psychological Science, 20*(6), 729–739. https://doi.org/10.1111/j.1467-9280.2009.02365.x

Jones, S., & Brandt, S. (2019). Do children really acquire dense neighbourhoods? *Journal of Child Language, 46*(6), 1260–1273. https://doi.org/10.1017/S0305000919000473

Kent, R. D., & Rountrey, C. (2020). What acoustic studies tell us about vowels in developing and disordered speech. *American Journal of Speech-Language Pathology, 29*(3), 1749–1778. https://doi.org/10.1044/2020_AJSLP-19-00178

Kern, S., & Gayraud, F. (2010). *IFDC.* Grenoble: Les éditions la cigale.

Khattab, G., & Al-Tamimi, J. (2013). Influence of geminate structure on early arabic templatic patterns. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic evidence* (pp. 374–414). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511980503.018

Laing, C. (2024). Phonological Systematicity in Development. Retrieved from osf.io/uzrsy

Luef, E. M. (2022). Growth algorithms in the phonological networks of second language learners: A replication of Siew and Vitevitch (2020a). *Journal of Experimental Psychology: General, 151*(12), e26–e44. https://doi.org/10.1037/xge0001248

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.1177/026565909200800211

Mak, M. H. C., & Twitchell, H. (2020). Evidence for preferential attachment: Words that are more well connected in semantic networks are better at acquiring new links in paired-associate learning. *Psychonomic Bulletin & Review, 27*(5), 1059–1069. https://doi.org/10.3758/s13423-020-01773-0

McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical development. *Journal of Speech, Language, and Hearing Research, 44*(3), 670–684. https://doi.org/10.1044/1092-4388(2001/054)

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100

Mitchell, L., Tsui, R. K. Y., & Byers-Heinlein, K. (2022). *Cognates are advantaged in early bilingual expressive vocabulary development.* PsyArXiv. https://doi.org/10.31234/osf.io/daktp

Monaghan, P., Christiansen, M. H., Farmer, T. A., & Fitneva, S. A. (2010). Measures of phonological typicality. *The Mental Lexicon, 5*(3), 281–299. https://doi.org/10.1075/ml.5.3.02mon

R Core Team. (2020). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from

https://www.R-project.org/

Rij, J. van, Wieling, M., Baayen, R. H., & Rijn, H. van. (2022). *itsadug: Interpreting time series and autocorrelated data using GAMMs.*

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Seidenberg, M. S., Harm, M. W., & Seidenberg, M. S. (1999). Dyslexia: Insights from connectionist models. *Unpublished Doctoral Dissertation, University of Southern California, Los*, 491–528.

Siew, C. S. Q., & Vitevitch, M. S. (2020). An investigation of network growth principles in the phonological language network. *Journal of Experimental Psychology: General.* https://doi.org/10.1037/xge0000876

Sóskuthy, M. (2017). *Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction.* arXiv. Retrieved from http://arxiv.org/abs/1703.05339

Stamer, M. K., & Vitevitch, M. S. (2012). Phonological similarity influences word learning in adults learning Spanish as a foreign language. *Bilingualism: Language and Cognition*, *15*(3), 490–502. https://doi.org/10.1017/S1366728911000216

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41–78. https://doi.org/10.1207/s15516709cog2901_3

Stoel-Gammon, C., & Cooper, J. A. (1984). Patterns of early lexical and phonological development. *Journal of Child Language*, *11*(2), 247–271. https://doi.org/10.1017/S0305000900005766

Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, *25*(2),

201–221. https://doi.org/10.1017/S0142716404001109

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating Phonotactic Probability and Neighborhood Density in Adult Word Learning. *Journal of Speech, Language, and Hearing Research : JSLHR*, *49*(6), 1175–1192. https://doi.org/10.1044/1092-4388(2006/085)

Szreder, M. (2013). The acquisition of consonant clusters in polish: A case study. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic evidence* (pp. 343–361). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511980503.016

Thelen, E., & Smith, L. (1996). *A dynamic systems approach to the development of cognition and action.* Boston, MA: MIT Press.

Vihman, M. M. (2014). *Phonological development: The first two years* (2nd ed.). Malden, MA: Wiley-Blackwell.

Vihman, M. M. (2016). Prosodic structures and templates in bilingual phonological development. *Bilingualism: Language and Cognition*, *19*(1), 69–88. https://doi.org/10.1017/S1366728914000790

Vihman, M. M. (2019). *Phonological templates in development.* Oxford, UK: Oxford University Press.

Vihman, M. M., Kay, E., Boysson-Bardies, B. de, Durand, C., & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to 1-year-old children. *Developmental Psychology*, *30*(5), 651–662.

Vihman, M. M., & Keren-Portnoy, T. (2013). The emergence of phonology: Whole-word approaches, cross-linguistic evidence. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic evidence.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511980503.002

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of Memory and Language*, *40*(3), 374–408. https://doi.org/10.1006/jmla.1998.2618

Waterson, N. (1971). Child phonology : A prosodic view. *Journal of Linguistics*, *7*(2), 179–211. https://doi.org/10.1017/S0022226700002917

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, *73*(1), 3–36.

\end{longtable}

Table 1

Age (months) at first session, number of sessions and number of distinct word types and tokens produced by each child in the dataset. Means and SDs for each corpus are shown in bold.

Speaker	Corpus	Min. age	n Sessions	Types	n Tokens
Anais	French	12	17	283	7169
Marie	French	12	14	258	5677
Nathan	French	12	17	162	4814
Tim	French	11	17	460	11489
Mean	**French**	**12**	**16**	**291**	**7287**
SD	**French**	**0**	**2**	**124**	**2965**
Alex	English	16	14	272	5253
Lily	English	13	16	456	8221
Naima	English	11	19	550	8107
Violet	English	14	14	385	6604
William	English	16	13	270	2888
Mean	**English**	**14**	**15**	**387**	**6215**
SD	**English**	**2**	**2**	**121**	**2222**
Mean	**All**	**13**	**16**	**344**	**6691**
SD	**All**	**2**	**2**	**125**	**2467**

Table 2

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

| Model		Actual			Target	
	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 3

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.

| Effect	Actual					Target				
	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]

Table 4

Outputs from nested model comparisons of GAMMs testing the effect of age on INT and EXT values in Actual and Target data (Models 1 and 2), and the effect of Data type on INT values (Model 3). Model comparisons compared full models against those without parametric and smooth terms that included the variable being tested.

|	Model	Actual			Target		
		Df	Chi Sq	p	Df	Chi Sq	p
1	INT:Age	14.000	17.886	0.001	14.000	33.594	<.001
2	EXT:Age	14.000	6.088	0.592	14.000	10.084	0.125
3	INT:Data type	7.000	496.560	<.001			
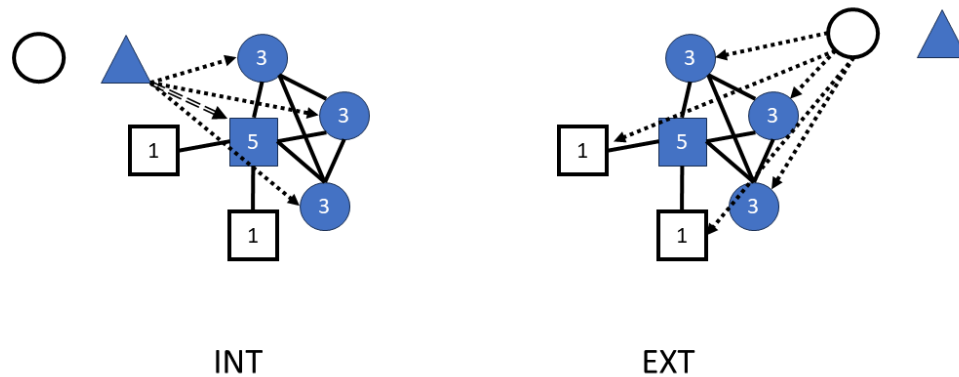
Figure 1. Visualisation of INT and EXT models of network growth. Shapes represent nodes in the network and filled lines represent edges between nodes. Numbers show the degree of each existing node in the network. The two images demonstrate the likelihood of two new nodes - a filled triangle or an open circle - being added to the network under conditions of INT- and EXT-like network growth. In each case, the node that would be acquired is added to the network, and new edges are shown with dashed arrows. The double-dashed arrow in the INT model shows the new edge formed with the most highly-connected node in the existing network.
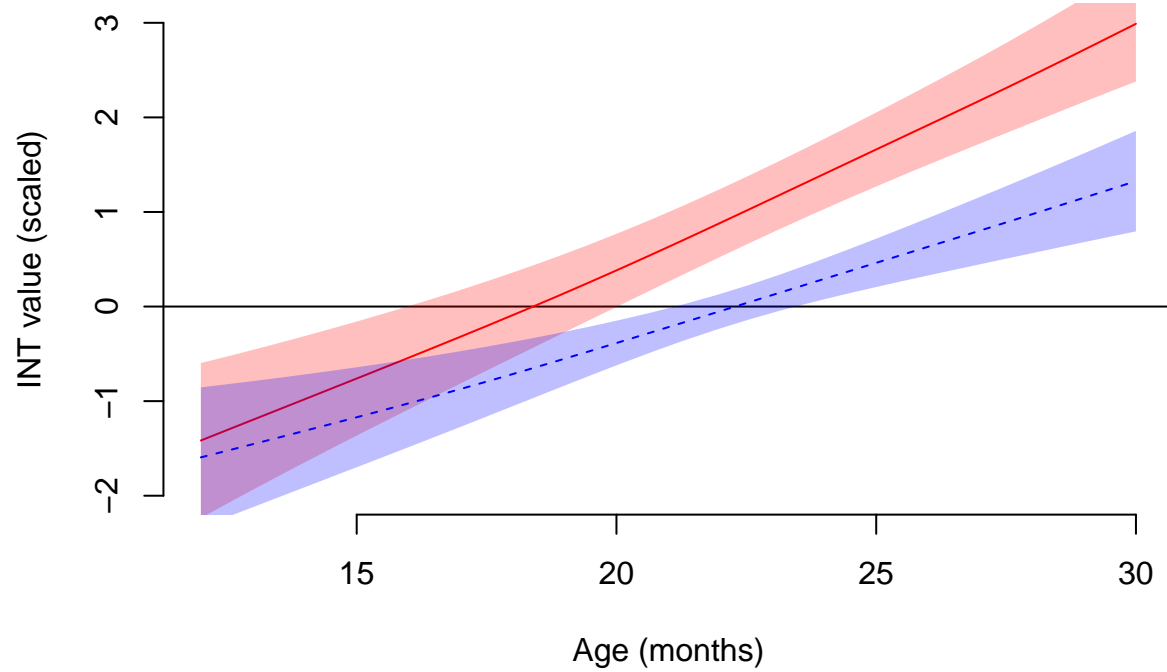
Figure 2. INT values over time in Actual and Target data. Red filled line represents Actual values, blue dashed line represents Target values; coloured bands represent 95% CIs.
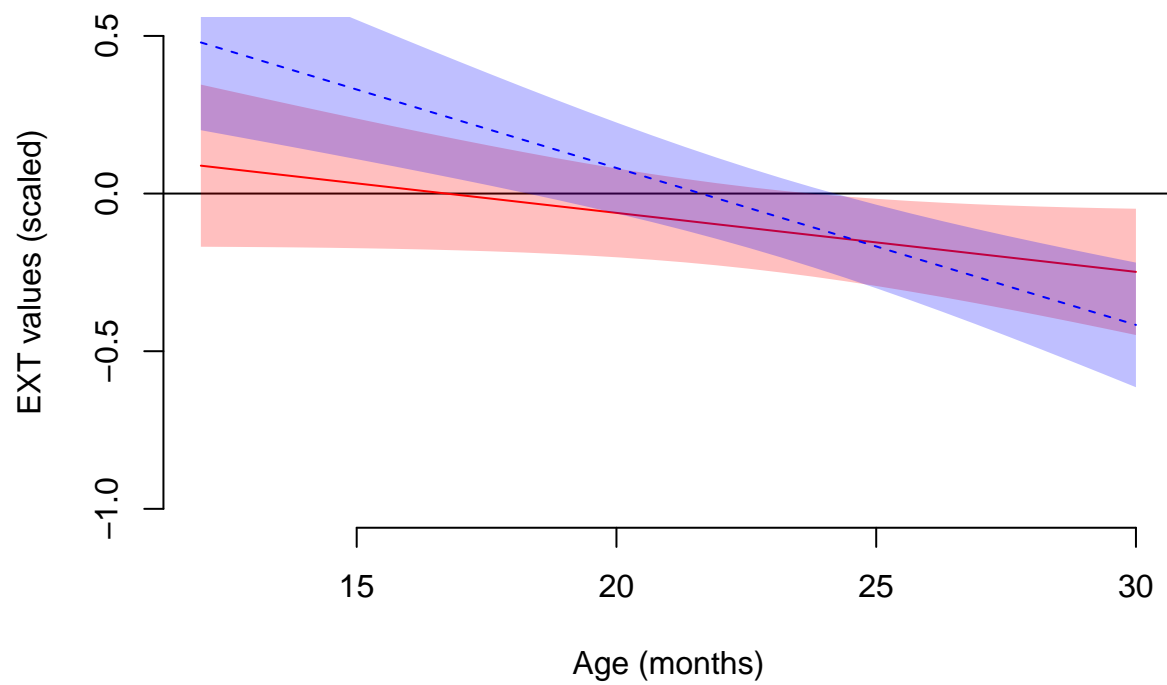
Figure 3. EXT values over time in Actual and Target data. Red filled line represents Actual values, blue dashed line represents Target values; coloured bands represent 95% CIs. Both smooths are shown here for exploratory purposes.
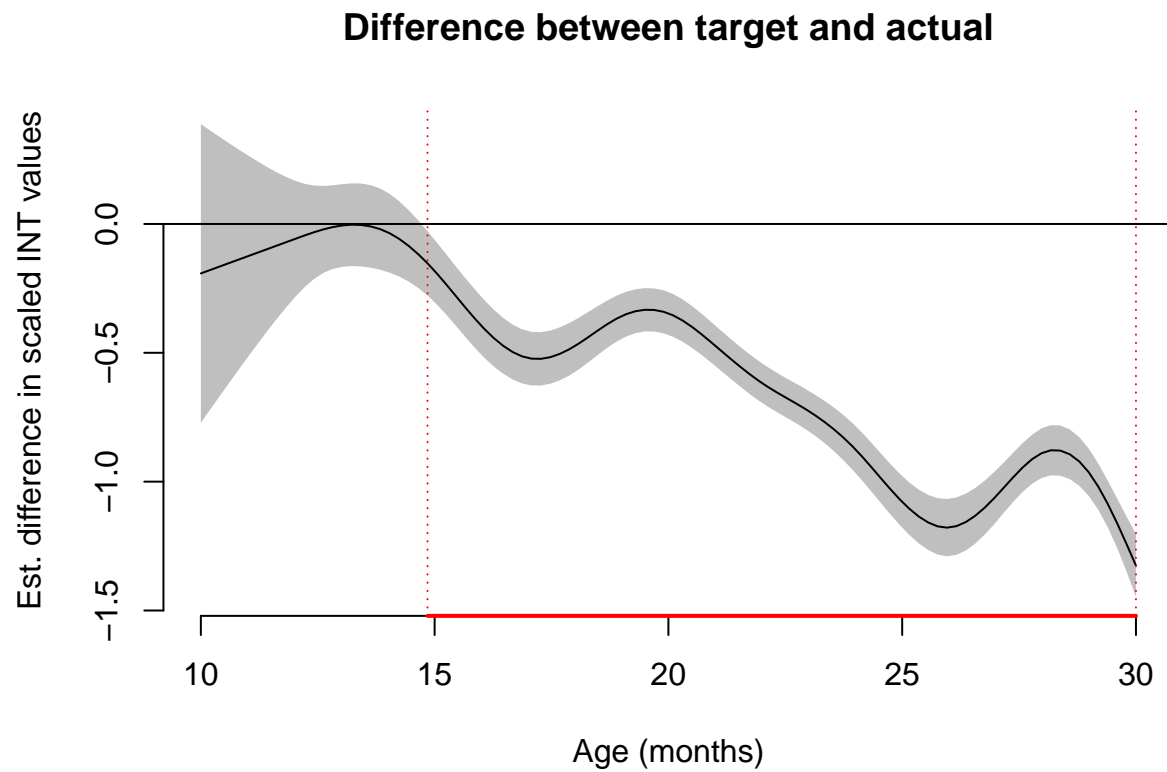
Figure 4. Difference smooth plot showing difference between scaled INT values in Actual vs. Target forms from the GAMM model specified above. Shaded area shows 95% confidence intervals, red line along x-axis indicates months in which the difference between Actual and Target forms was significant.
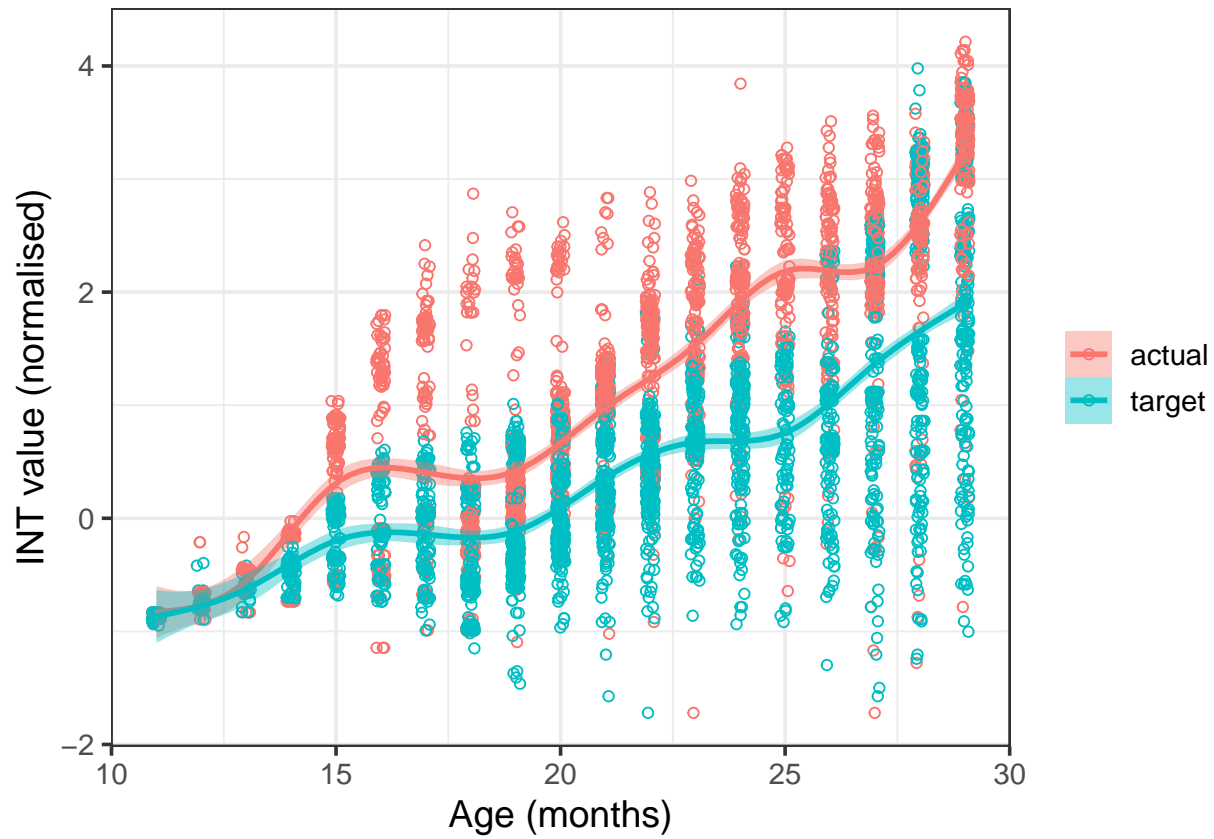
Figure 5. Smooth plot showing scaled INT values in Actual vs. Target forms. Shaded areas show 95% confidence intervals, lines indicate mean trajectories over time, coloured circles represent individual datapoints, jittered for visual clarity.