

How children explore the phonological network in child-directed speech: A survival analysis of  
children's first word productions

Matthew T. Carlson<sup>a, b, \*</sup>, Morgan Sonderegger<sup>c, d, e</sup>, and Max Bane<sup>d</sup>

\*Correspondence concerning this article should be sent to Matthew T Carlson.

<sup>a</sup> (Present address)

The Pennsylvania State University  
Department of Spanish, Italian, and  
Portuguese  
231 Burrowes Bldg.  
University Park, PA 16802  
USA

[mtc173@psu.edu](mailto:mtc173@psu.edu)

1-814-865-0104

<sup>b</sup> University of Chicago

Department of Psychology  
5848 S. University Ave.  
Chicago, IL 60637  
USA

<sup>c</sup> (Present address)

McGill University  
Department of Linguistics  
1085 Dr. Penfield, Room 111  
Montreal, Quebec  
H3A 1A7  
Canada

[morgan.sonderegger@mcgill.ca](mailto:morgan.sonderegger@mcgill.ca)

<sup>d</sup> University of Chicago

Department of Linguistics  
1010 E. 59<sup>th</sup> St.  
Chicago, IL 60637  
USA

<sup>e</sup> University of Chicago

Department of Computer Science  
1100 E. 58<sup>th</sup> St.  
Chicago, IL 60637  
USA

## Abstract

We explored how phonological network structure influences the age of words' first appearance in children's (14-50 months) speech, using a large, longitudinal corpus of spontaneous child-caregiver interactions. We represent the caregiver lexicon as a network in which each word is connected to all of its phonological neighbors, and consider both words' local neighborhood density (*degree*), and also their embeddedness among interconnected neighborhoods (*clustering coefficient* and *coreness*). The larger-scale structure reflected in the latter two measures is implicated in current theories of lexical development and processing, but its role in lexical development has not yet been explored. Multilevel discrete-time survival analysis revealed that children are more likely to produce new words whose network properties support lexical access for production: high degree, but low clustering coefficient and coreness. These effects appear to be strongest at earlier ages and largely absent from 30 months on. These results suggest that both a word's local connectivity in the lexicon and its position in the lexicon as a whole influences when it is learned, and they underscore how general lexical processing mechanisms contribute to productive vocabulary development.

*Keywords:* Phonological development; phonological networks; vocabulary growth; network science; neighborhood density; clustering coefficient; coreness; survival analysis

How children explore the phonological network in child-directed speech: A survival analysis of children's first word productions

Why do children systematically produce some words at an earlier age than other words?

What biases guide word learning, and how might these biases change as the child develops?

Researchers have identified a variety of word properties that influence acquisition, including semantic, morphosyntactic, and formal properties (e.g. Gentner & Boroditsky, 2001; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Hills, Maouene, Riordan, & Smith, 2010; Hills, Maouene, Maouene, Sheya, & Smith, 2009; Stevens, Yang, Trueswell, & Gleitman, 2012; Steyvers & Tenenbaum, 2005; Stoel-Gammon, 2011; Vihman & Velleman, 2000). Some of these are properties of the word itself, and others concern relationships among words on semantic or formal dimensions. In the latter group, much attention has focused on how the phonological similarity of single words to other words in the rest of the lexicon influences ease of acquisition. The most common and long-standing operational definition of phonological similarity involves *phonological neighbors*, words that differ by the addition, deletion, or substitution of one phoneme (Landauer & Streeter, 1973). The number of neighbors that a target word has, based on this definition, is known as the target's (*phonological*) *neighborhood density*.

The focus on neighborhood density has yielded important insight into the role of phonological similarity in language development as well as lexical processing (Bernstein Ratner, Newman, & Storkel, 2009; Bernstein Ratner et al., 2009; Charles-Luce & Luce, 1990; Coady & Aslin, 2003; Garlock, Walley, & Metsala, 2001; Luce & Large, 2001; Metsala, 1997; R. S. Newman & German, 2002; Stoel-Gammon, 2011; Storkel, 2004; Swingley & Aslin, 2002; Vitevitch, Luce, Pisoni, & Auer, 1999; Vitevitch & Luce, 1998, 1999; Vitevitch, 2002), and neighborhood density is a central concept in prominent theories in these domains (Dell & Gordon, 2003; Luce & Pisoni, 1998; Metsala & Walley, 1998; Walley, 1993). However, the results of neighborhood density research are complex, and researchers have long grappled with a sense that neighborhood density does not capture

important aspects of phonological similarity between words in the lexicon (e.g. Bailey & Hahn, 2005; Mathey, Robert, & Zagar, 2004; Mathey & Zagar, 2000; Yarkoni, Balota, & Yap, 2008; Zamuner, 2009).

Some of these problems stem from a formal consequence of the traditional definition of phonological neighbor: a word's neighbors usually have neighbors of their own, which in turn have other neighbors, and so on. The definition of phonological neighbors thus implicitly defines a representation of an entire lexicon in which each word is connected to all of its neighbors, a *phonological network* (Arbesman, Strogatz, & Vitevitch, 2010a, 2010b; Vitevitch, 2008). In the present study, we apply this definition over a large corpus of child-directed speech to construct a phonological network that approximates the lexical exposure of American English speaking preschoolers.

The structure of complex networks such as this may be quantified both at and beyond the scale of local neighborhoods (M. E. Newman, 2003), potentially shedding light on the relationship between when children learn a word and how the word is embedded in the phonological network. We focus on three common network-theoretic measures, all defined at the level of individual words: traditional neighborhood density (henceforth referred to using the equivalent network-theoretic term *degree*), *clustering coefficient* and *coreness* (all defined below). The impact of a word's degree on when it is learned has been extensively studied in previous work, and clustering coefficient and coreness are particularly well-suited to measuring the kind of larger-scale network structure implicit in current theories of lexical and phonological development (cf. recent attention to clustering coefficient in research on adult lexical processing Chan & Vitevitch, 2009, 2010; Yates, 2013). We will use the term *phonological network properties* to denote degree, clustering coefficient, and coreness, which together provide a richer description of each word's phonological relationship to the rest of the lexicon than is available from degree alone.

We assess the power of these phonological network properties for predicting when new words enter children's productive lexicons, using a large longitudinal corpus of spontaneous child-caregiver interaction spanning child ages 14-50 months (Rowe & Goldin-Meadow, 2009; Rowe, Raudenbush, & Goldin-Meadow, 2012). Using longitudinal data allows us to examine both how phonological network structure affects word learning, and how its role changes as children develop. We do so using *survival analysis*, a statistical technique for modeling the time elapsed prior to some event, here the first time a target word is observed in the spontaneous speech of a given child (Barber, Murphy, Axinn, & Maples, 2000; Reardon, Brennan, & Buka, 2002; Singer & Willett, 1991, 2003). Survival analysis also allows us to control for a set of covariates known to impact word learning: frequency, length, syntactic category, phonotactic probability, child gender, and quantity of caregiver speech input.

Our study thus focuses on two main questions about children's productive vocabulary growth:

1. Do children produce some words earlier than others based on both local and larger-scale phonological network properties in child-directed speech? If so, what are the directions of the effects of degree, clustering coefficient, and coreness?
2. Does children's sensitivity to these properties change over time, and if so, how?

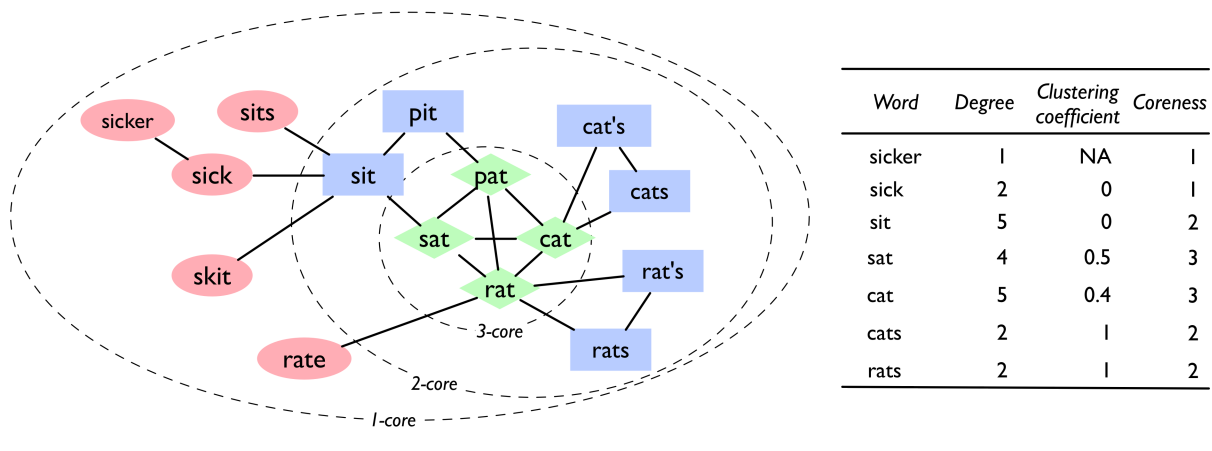
We begin with an overview of aspects of network science which are relevant for our investigation, then contextualize the network-theoretic approach within the literature on neighborhood density and review recent efforts to apply it to lexical organization, processing, and acquisition. We then describe our data, analytic strategy, and results.

## **Background**

### **The Lexicon as Phonological Network**

A complex network consists of a set of *nodes* and a set of *edges* linking pairs of nodes based on some *edge condition* (M. E. Newman, 2003). In a phonological network of the kind examined here

(see the network fragment in Figure 1), the nodes are the words in the lexicon, and the most common edge condition is the traditional definition of a phonological neighbor (Landauer & Streeter, 1973): two words are linked if they differ by the addition, subtraction, or substitution of a single phoneme (Arbesman et al., 2010a, 2010b; Vitevitch, 2008), based on words' adult-like segmental composition. While this definition neglects the role of features, suprasegmentals, the position of the edit, and so on (e.g. Bailey & Hahn, 2005; Mathey et al., 2004; Mathey & Zagar, 2000; Yarkoni et al., 2008; Zamuner, 2009), considering an entire network substantially enriches the concept of phonological neighborhoods with the capacity to quantify larger-scale structure.



**Figure 1.** Left: a fragment of the English phonological network. Each node represents an orthographic word; edges (solid lines) connect words which are phonological neighbors. Dashed lines show the  $k$ -core decomposition of the fragment. Right: values of phonological network properties investigated in this study (degree, clustering coefficient, coreness), for a subset of words in this network. Clustering coefficient is undefined for nodes with degree  $< 2$ , such as *sicker*, because there are no possible triangles including these nodes.

The most local property of nodes in a network is their *degree*, which in the present case corresponds to neighborhood density: the number of edges connected to the word. *Clustering coefficient* expands the focus to take in properties of a target word's neighbors, and is defined as the proportion of all pairs of a target's neighbors that are neighbors of each other. For example, in Figure 1, *sat* has four neighbors, *sit*, *pat*, *cat*, and *rat*. There are thus six possible pairings of its

neighbors: *sit-pat*, *sit-cat*, *sit-rat*, *pat-cat*, *pat-rat*, and *cat-rat*, but only the last three of these combinations represent pairs of neighbors. The clustering coefficient of *sat* in this network fragment is thus  $3/6$ , or 0.5.

The most global phonological network property of individual words that we examine here is *coreness*, which quantifies a word's embeddedness in the network (Alvarez-Hamelin, Dall'Asta, Barrat, & Vespignani, 2005, 2008; Dorogovtsev, Goltsev, & Mendes, 2006). It is the most global because it is calculated based on the structure of the entire network: a node (word) has coreness  $k$  if it both remains in the network after recursively pruning all nodes with degree  $< k$ , and is removed after recursively pruning all nodes with degree  $\leq k$ . Note that a word's coreness may not exceed its degree, but it may be smaller. For example, while *sick*, in Figure 1, has two neighbors, it resides in the first core because after removing *sicker*, which has degree 1, *sick* has only one remaining neighbor and is therefore also pruned. Words with higher coreness belong to increasingly cohesive subgroups of the network, which are intuitively "deeper" in the network, and words with lower coreness belong to less cohesive subgroups of the network, which are intuitively more "peripheral".

### **Phonological Network Structure and Acquiring a Lexicon**

Coreness and clustering coefficient quantify aspects of the interrelatedness of local neighborhoods, making them well-suited to our primary goal of investigating the role of larger-scale phonological similarity structure in word learning. Current theory attributes degree effects in word learning to the ways in which children represent and process phonologically similar words, and we will argue below that a role for larger-scale network structure is also implicit in both of these mechanisms. This leads us to expect significant effects of clustering coefficient and coreness, in addition to degree, on children's word learning.

However, the current literature does not provide a clear indication of what the direction of these effects should be, and in fact there is little reason to expect all three phonological network properties to relate to word learning in the same way. In general, phonological similarity has been

thought of as both a challenge (because similar words are more confusable; Ferguson & Farwell, 1975; Hallé & de Boysson-Bardies, 1996; Storkel & Lee, 2011; Vihman & Velleman, 2000; Vihman, 1996), and an opportunity (because words with many neighbors tend to be composed of frequent phonological material; Lindblom, 1992; Menn, 1978; or because neighbors can support novel words' representations in working memory; Storkel & Lee, 2011).

Following the first of these lines of reasoning, the challenge of representing phonologically similar words in memory is thought to drive the emergence of segmental detail in children's phonological representations, an idea most extensively developed in the Lexical Restructuring Hypothesis (Beckman, Munson, & Edwards, 2007; Metsala & Walley, 1998; R. S. Newman, 2008; Stoel-Gammon, 2011; Walley, 1993). Since words with many neighbors can only be distinguished from those neighbors through the representation of rich segmental detail, but less well specified representations are sufficient for words with few neighbors, detail (or the ability to use detail efficiently, R. S. Newman, 2008) emerges first in children's representations of words with high degree (Garlock et al., 2001).

The second line of reasoning is apparent in a more articulated view of the development of phonological representations recently proposed by Storkel and Lee (2011), in which children must first identify new words requiring a new lexical entry (*triggering*), then develop a specified representation for the word (*configuration*) and integrate that representation with the rest of the lexicon (*engagement*). Under this view, confusability with many neighbors may inhibit triggering by making it harder to distinguish a new word from other, existing representations, but neighbors may be more helpful in the later stages of word learning by facilitating the processing of new words or helping to maintain them in working memory longer. Other ways of dividing up the process of word learning may be possible (e.g. Werker & Curtin, 2005), but Storkel and Lee's account points out how the similarity of a word with its neighbors may result in confusability under some circumstances,



and support under others. Crucially, a role of the interconnectedness of neighborhoods is implicit in both of these mechanisms.

Regarding the demands that neighbors place on representation, we might at first glance consider densely interconnected neighborhoods to simply increase the general level of confusability across a wider portion of the network, making acquisition, or triggering, more difficult where degree, clustering coefficient, and coreness are all high. On the other hand, the confusability of a new word with its immediate phonological neighbors may well depend on the pressure to represent those neighbors in detail, implicating both clustering coefficient and coreness. Where clustering coefficient is high, a target word's neighbors tend to differ from the target in few segmental positions and the neighbors' relationships with each other thus highlight precisely the detail that will help distinguish the target. Where coreness is high, many dense neighborhoods are interconnected, supporting the emergence of detailed representations across a larger area of the lexicon. High clustering coefficient and coreness may thus in a sense pave the way for the addition of new words, since better specified representations of their neighbors should reduce confusability.

Regarding lexical processing, a role for clustering coefficient and coreness is implicit in the basic idea of spreading activation (e.g. Luce & Pisoni, 1998), because there is nothing to prevent activation from flowing beyond a target word's neighbors to their own neighbors and beyond, apart from gradual decay at greater distances (Chan & Vitevitch, 2009). Chan and Vitevitch suggested that the interconnectedness of neighbor relationships among the words surrounding a target, which they measured using clustering coefficient, can affect the gradient of activation between the target and its competitors (see also Chen & Mirman, 2012) by allowing activation to either dissipate farther out in the network (low clustering coefficient) or become pooled among competitors (high clustering coefficient). This accords with the inhibitory effects of clustering coefficient on adult performance in both word recognition (Altieri, Gruenenfelder, & Pisoni, 2010; Chan & Vitevitch, 2009; Yates, 2013) and production (Chan & Vitevitch, 2010) tasks, and a similar scenario is

expected for coreness, where the interconnectedness of neighborhoods provides paths whereby activation may flow among competitors, boosting their activation relative to the target. If working memory and other lexical processing mechanisms play a substantial role in the development of new phonological representations, we would therefore expect inhibitory effects of clustering coefficient and coreness on acquisition, in contrast to the facilitatory effects reported for degree (Storkel & Lee, 2011; cf. the facilitatory effects of degree on speech production, Vitevitch, 2002; Vitevitch & Sommers, 2003).

The emerging theoretical picture is thus complex. The initial recognition of new words may be inhibited by high degree, clustering coefficient, and coreness because of their high confusability among known competitor words, although the effects of clustering coefficient and coreness could instead turn out to facilitate word learning, because high values for these properties are expected to support more detailed representations of a new target's direct competitors, helping to distinguish the target. If, however, general processes involved with lexical access substantially mediate lexical development, at least following children's first encounter with a word, then we might expect facilitatory effects of degree, but inhibitory effects of clustering coefficient and coreness. This latter possibility is particularly compatible with a view of lexical development as a process of learning to produce words that the child has already begun to recognize (Stokes, 2014).<sup>1</sup>

The existing evidence, focused primarily on degree, does not unambiguously support any of these possibilities. Prior investigations of this issue have relied on two major strategies. The first explores differences in the overall composition of child and adult lexicons as evidence for children's biases in word learning. The second examines whether individual words or non-words are more or less difficult to acquire (in naturalistic or laboratory contexts) based on the presence of neighbors.

---

<sup>1</sup> Note that this only requires that children have begun to develop phonological representations of the neighbors; they need not be able to produce them yet, nor even necessarily comprehend them fully.

There is less empirical evidence concerning our second research question, how children's sensitivity to phonological similarity changes over time. From the existing literature on this issue, summarized below, the expectation that sensitivity to phonological similarity *does* change seems clear, but there are several possibilities for what type of change should take place.

**Studies of phonological similarity structure across the entire lexicon.** Focusing on the overall composition of the lexicon, Charles-Luce & Luce (1990, 1995) predicted that, if representing confusable words is a challenge for children, then the early lexicon should be characterized by sparser neighborhoods than the adult lexicon. Comparing the lexicons of 5 and 7 year-olds with an adult lexicon (a dictionary), they found support for this view, but the fact that most words in the child lexicon had at least some neighbors, and some had many, appears to indicate substantial sensitivity to detail (Dollaghan, 1994). Moreover, Coady and Aslin (2003) argued that since children's lexicons are smaller than adults', it might be more meaningful to measure the proportion of the lexicon to which a target word was similar, rather than the raw number of neighbors. Using this metric, they obtained the opposite result: children's lexical neighborhoods (age 42 months) were denser, not sparser, than adults'.

In network-theoretic terms, these studies compared the *degree distributions* of phonological networks, which is one of several measures of global network structure. For example, in (adult) phonological networks in a variety of languages (Arbesman et al., 2010a, 2010b), most of the lexicon resides in a single interconnected "giant component", and the rest resides in many smaller "islands." In addition, the degree distribution tends to be fit well by a truncated power law (Arbesman et al., 2010b), the *average shortest distance* between any randomly selected pair of nodes is small, the average clustering coefficient is high, and edges tend to connect nodes with similar degrees, a property known as *assortative mixing*. Networks (linguistic or otherwise) with

these properties are often said to exhibit *small-world* structure,<sup>2</sup> and have been argued to be easily searchable and to exhibit stable structure in the face of missing nodes (Amaral, Scala, Barthélemy, & Stanley, 2000; Kleinberg, 2000; Watts & Strogatz, 1998).

Carlson, Bane, and Sonderegger (2011) used several of these metrics (including degree distribution, average degree, average shortest path, assortativity, and mean clustering coefficient) to compare lexicons derived from corpora of child-, child-directed, and adult-directed speech. The child- and child-directed speech lexicons, approximating the productive vocabularies and cumulative input of 4 year-old children, exhibited values more indicative of small-world networks than the adult lexicon, suggesting that the growth process underlying child phonological networks may be biased to favor searchability and stability.

One interpretation of the findings of Carlson et al. (2011) and Coady and Aslin (2003) is that children acquire words in a way that favors a “denser” lexicon, which in the context of the current study would lead us to expect facilitatory effects of all three phonological network properties. However, small-world structure per se does not necessarily indicate anything about the biases that guide children’s piecemeal acquisition of the lexicon. Children’s lexicons may simply replicate the properties of their input, as suggested by the highly similar network properties found in these studies between the child- and child-directed speech networks, despite the latter being much larger. Moreover, while small-world structure can be a sign of certain kinds of growth processes in other types of networks (Albert & Barabási, 2002), a peculiarity of phonological networks necessarily yields small-world structure regardless of the underlying growth process (Gruenenfelder & Pisoni, 2009).<sup>3</sup>

---

<sup>2</sup> More precisely, small-world networks are defined as having a low average shortest path length, and a high clustering coefficient, compared to random networks.

<sup>3</sup> Unlike many other kinds of networks, the nodes in a phonological network (words) comprise strings of lower-level units from a finite inventory (phonemes). Gruenenfelder and Pisoni constructed a set of such networks based on lexicons of randomly generated non-words, and found that the resulting phonological networks had the same small-world properties noted by Vitevitch (2008) for the English phonological network.

Thus, while studies evaluating the overall composition of the phonological network suggest that children may favor a “denser” lexicon in some sense (higher degree, and properties more consistent with small-world networks), the sensitivity of these findings to methodological decisions, as well as issues of how to interpret small-world structure, make this suggestion tentative at best.

**Phonological similarity and the learning of individual words.** The research on effects of phonological similarity on the learning of individual words has not yet addressed larger-scale phonological network properties such as clustering coefficient and coreness, and for degree the emerging picture is complex. Experimental studies using both non-words and words provide ample evidence that children find highly similar words to be more challenging to learn than less similar words across the first 5 years of life (Pater, Stager, & Werker, 2004; Stager & Werker, 1997; Swingley & Aslin, 2007; but see Fennell & Werker, 2003; McKean, Letts, & Howard, 2013; Werker & Curtin, 2005; Yoshida, Fennell, Swingley, & Werker, 2009). However, facilitatory effects of phonological similarity emerge under certain conditions. For example, while multiple repetitions of phonological neighbors led to reduced non-word learning in 17 month olds, exposure to only a few repetitions of the neighbors led to enhanced learning (Hollich, Jusczyk, & Luce, 2002). Similarly, hearing known neighbors boosted 2 year-olds’ recognition (Merriman & Marazita, 1995), and 4 year-olds’ production (Demke, Graham, & Siakaluk, 2002), of novel nouns. Finally, Storkel and Lee (2011) found better learning of low-degree than high-degree non-words on an immediate posttest, but this pattern was reversed on a delayed posttest a week later.

Thus, while children may have difficulty with direct conflict between similar words, phonological neighbors may support word learning as long as competition is not too great (e.g. when the neighbors are already known to the child or when they are not too salient), or they may

---

support some underlying processes of word learning (e.g. configuration and engagement), but not others (triggering).

There have been fewer investigations of the role of phonological similarity in word learning with naturalistic data. The analyses of the overall composition of the child lexicon reviewed above (Carlson et al., 2011; Charles-Luce & Luce, 1990, 1995; Coady & Aslin, 2003), used naturalistic data, but we have already reviewed the difficulties in interpreting their results. Storkel (2009, see also 2004) analyzed the timing of nouns' appearance in children's (aged 16-30 months) productive lexicons, based on the proportion of children of each age reported by their parents to have produced the nouns on the MacArthur-Bates CDI (Fenson et al., 1994). Short, high degree words were known by more children and acquired at an earlier age (but see Stokes, 2010; Stokes, Kern, & Dos Santos, 2012), but the strength of the effect decreased steadily after age 2. This finding does not conclusively demonstrate a degree effect, as Storkel used a composite score of length and degree, but it yields some support for predicting a positive effect of degree here. There is also evidence for substantial variability in children's sensitivity to degree (Maekawa & Storkel, 2006). To our knowledge, the current study is the first to examine the effects of clustering coefficient and coreness on word learning.

**Developmental change in sensitivity to phonological similarity.** Storkel's (2009) finding supports the expectation of developmental change in the effects of phonological similarity, as do differences in the results of experimental studies using different age ranges, cited in the previous subsection. The early effects of phonological similarity may be positive or negative, depending on various factors, but in either case the effects seem to attenuate with age. Developmentally graded effects are also expected under the PRIMIR model of infant speech processing (Werker & Curtin, 2005).

Attenuation is also supported by the idea described above that children's ability to use phonological detail is driven by the need to distinguish similar words in memory (Beckman et al.,

2007; Ferguson & Farwell, 1975; Hallé & de Boysson-Bardies, 1996; Vihman, 1996). Since phonological detail is expected to emerge first in denser areas of the lexicon, and then spread to sparser areas, the greatest difference between dense and sparse areas should occur early and attenuate as words in the sparser areas ‘catch up’ to those in denser areas. This should be true for both local (degree) and larger scale (clustering coefficient and coreness) phonological network properties, based on the reasoning above. However, the developmental attenuation of effects may or may not be apparent in the age range examined here (14-50 months), given that lexical restructuring is likely to continue well into later childhood (Metsala & Walley, 1998; Walley, 1993).

The effects of phonological network structure on working memory or other lexical access processes, however, persist in adulthood, as evidenced by the extensive literature on degree and the growing literature on clustering coefficient (Altieri et al., 2010; Chan & Vitevitch, 2009, 2010; Luce & Large, 2001; Vitevitch et al., 1999; Vitevitch & Luce, 1998, 1999; Vitevitch & Sommers, 2003; Vitevitch, 2002; Yates, 2013). Thus, phonological network structure may continue to impact word learning in adults (Storkel, Armbruster, & Hogan, 2006). To the extent that this mechanism is implicated in children’s productive vocabulary growth as well (Storkel & Lee, 2011), we might expect it to be fairly stable developmentally. The question is whether we will be able to detect any effects of phonological network structure on processing in spontaneous speech data. When children’s lexicons are small, the effects of ease of processing on productive vocabulary may be apparent (Storkel, 2009), but later on children’s word choice is more likely to reflect their communicative needs than their lexical knowledge or ease of lexical processing (Rowland, Fletcher, & Freudenthal, 2008; Tomasello & Stahl, 2004). Thus, an apparent attenuation of phonological network effects on productive vocabulary may not indicate a change in the underlying mechanism.

In the present study we use the Chicago corpus (Rowe & Goldin-Meadow, 2009), a large longitudinal corpus of spontaneous child-caregiver speech, to test our expectation of developmentally graded effects of phonological network properties on the timing of words’ entry

into children's productive vocabularies. The caregiver speech samples offer an approximation of children's lexical experience, allowing us to measure words' network and other properties. The children's longitudinal samples allow us to observe the age at which words enter children's speech for the first time, forming the basis for our survival analysis. By analyzing how children sample words from the available input, we are also able to test whether children simply replicate the network properties of their input lexicon, or actively favor words with particular properties.

### **Materials and Methods**

#### **Data**

The Chicago corpus consists of speech samples from 64 child-caregiver dyads who were observed during everyday interaction in the home for 90 minutes every four months, from child age 14 to 58 months (Rowe & Goldin-Meadow, 2009). Caregivers were asked to do whatever they would normally do during the same time interval. In the present study, we use only the 10 samples up to age 50 months. Two of the dyads were excluded due to missing data. In addition, if a child missed a recording session (to keep the sample durations relatively uniform, sessions that were cut short at less than 80 minutes were also considered "missed", for our purposes), that child's data from subsequent sessions were removed, or "right-censored", in the terminology of survival analysis. This is because what is being modeled in our discrete-time survival analysis is the age at the session of a child's first production of a word (see the Analysis subsection, below). Thus, if a child missed a recording session, data points from subsequent sessions are not informative, because her first productions of words could have occurred during the missed session. The data from nine children were censored at ages varying from 18 months to 42 months. The remaining 53 children were present for all 10 sessions.

The data for our study thus encompass 569 sessions from 62 dyads. 30 of the 62 children were female, and in most dyads the primary caregiver was the mother (in one dyad it was the father, and in five dyads both parents participated as primary caregivers). All children were being



raised monolingually in English, and the sampling of families was designed to approximately span the range of socioeconomic status, assessed using a composite of income and caregiver education, in the greater Chicago area. Several metrics of the size of the corpus are given in Table 1.

**Table 1. Size measures of the Chicago corpus (up to and including the 50 month samples)**

<i>Corpus</i>	<i>Word tokens</i>	<i>Utterances</i>	<i>Word types</i>
Child speech	815,139	320,053	8,366
Child-directed speech	1,954,556	573,379	14,890

The initial transcription included all words produced by the child or the primary caregiver, including onomatopoeia and interjections. The speech of other individuals who might have been present, and caregiver speech to individuals other than the child were not transcribed. Children were credited with the word attempted, even if pronunciation varied, as long as the word could be identified. Morphological errors (e.g. *runned*) were transcribed as produced.

**Phonological network in child-directed speech.** For the present study, the ambient lexicon was considered to be all orthographically unique words uttered by any caregiver in the Chicago corpus ( $n = 14890$ ), with the restrictions listed here. Part of speech was determined using the MOR tagger in CLAN (MacWhinney, 2000; Sagae, Davis, Lavie, MacWhinney, & Wintner, 2010). In the case that one orthographic form corresponded to words in more than one syntactic category, the more frequent part of speech was used, based on the output of the tagger. Phonological forms for each word were obtained from the CMU pronouncing dictionary (Carnegie Mellon Speech Group, 1993), and words not appearing in CMUdict were omitted ( $n = 1325$ ).<sup>4</sup> Inflected forms were

<sup>4</sup> Note that we do not assume that children produce the words in an adultlike way. The use of adultlike representations of words does, however, reflect the fact that children must contend with the adult forms in their environment.

retained, such that the lexicon consisted of the full set of phonological forms encountered by children in the ambient lexicon as defined here ( $n = 13565$ ).

This lexicon was used to construct a phonological network of the kind described above. Construction of the network, as well as measurement of the phonological network properties (degree, clustering coefficient, and coreness) for each word, were performed in Python using the NetworkX library (Hagberg, Schult, & Swart, 2008). Each word in the lexicon corresponded to one node in the network and two nodes were linked by an edge if the phonological edit distance between the nodes was 0 or 1. Note that homophones—words with different orthography, but the same phonological form—were thus counted as neighbors. Degree and coreness were calculated for all words, and clustering coefficient for words with degree of 2 or more, because clustering coefficient is not defined for words with less than two neighbors.

**Child word acquisition data.** The analysis of children's word acquisition presented below uses a subset of the words in the ambient lexicon, as defined above, chosen based on several criteria. The first and most important restriction was to consider only the 652 words which were produced at least once by at least 30 of the 62 children in the child speech samples of the Chicago corpus (all of which were present in the caregiver corpus as well).

This restriction was one of two ways in which we addressed an important difficulty with using spontaneous speech to gauge vocabulary knowledge. The appearance of words in spontaneous speech is strongly influenced by talkers' need to use those words during the observation (Ota & Green, 2013; Rowland et al., 2008; Tomasello & Stahl, 2004), and as a result, we are unlikely to observe a very low-frequency word in a given child's spontaneous speech, even if the child knows the word. Placing the cutoff at about half the sample of children resembles the common practice of using the age at which half of children have acquired a word as its age-of-acquisition (e.g. Goodman, Dale, & Li, 2008). Thus, the words included in this analysis would all have an age of acquisition at or before 50 months, the latest age recorded in the sample.

The second way we sought to capture effects of usage or communicative need on our chances of observing first productions was by including word frequency (in caregiver speech) as a predictor in the model. Any effects of frequency are thus not interpretable as effects on word acquisition, per se, but by controlling for this variable we aimed to make any effects of the phonological network properties more directly interpretable as reflecting acquisition rather than usage.<sup>5</sup>

In addition to variations in communicative need, a second difficulty in assessing vocabulary knowledge from spontaneous speech is the possibility that children may differ in how “talkative” they are, either in general, or across individual recording sessions. Our multilevel survival model accounts for this possibility to some extent with a by-child random intercept and by-child random slopes for the effect of age (see below for full model specification). Additional fluctuations might be captured using a fixed effect covariate (e.g. number of tokens produced by the child at each session), or with by-session intercepts nested within child. However, we do not expect talkativeness to alter the effects of phonological network properties, so we do not include these effects in order to reduce model complexity.

From this set of 652 words, we further excluded 101 words matching one or more of several criteria:

1. Communicators (e.g. *hey, okay*; 34 words) and onomatopoeic words (3 words), which tend to be more variable or to occur in reduced form in spontaneous speech.
2. All 63 words with degree < 2, for which clustering coefficient is undefined.<sup>6</sup>

---

<sup>5</sup> Additionally, we refit the model reported below using a much lower threshold, including all 1239 words that were acquired by at least 10 children (instead of 30) by the end of sampling, and obtained the same qualitative results, including the attenuation of phonological network effects by about 30 months. Since the effect of communicative need on the likelihood of observing utterances of known words is stronger for lower frequencies, this offers some reassurance that the effects reported below are not influenced by this feature of spontaneous speech.

<sup>6</sup> An alternate strategy would be to assign these words clustering coefficient = 0 because they have no neighbors that are neighbors of each other. We refit the model reported below using this strategy and obtained the same qualitative results.

3. Seven words consisting of one phoneme (e.g. *l*), for which the notion of neighborhood does not make intuitive sense (these seven words are all neighbors, despite sharing no phonological material).

Thus, our analysis considers a set of 551 words, listed in Appendix 2: 213 nouns, 140 verbs, 114 modifiers, and 84 closed-class words. Descriptive statistics for these words are given in Table 2, both for the untransformed variables, and the log-transformed versions included in the survival model as predictors.<sup>7 8</sup>

**Table 2. Descriptive statistics of word properties analyzed in the survival model. All properties are calculated based on the caregiver lexicon used in the phonological network. Frequency was measured using the caregiver portion of the Chicago corpus ( $n = 1,954,556$  tokens), from which the present data are taken.**

<i>Predictor</i>	<i>Mean</i>	<i>Median</i>	<i>SD</i>	<i>Range</i>
Word length	3.46	3	0.964	(1.5, 7)
log(word length)	0.930	0.811	0.373	(-0.288, 1.83)
Phonotactic probability	0.0202	$6.89 \cdot 10^{-5}$	0.0693	( $1.41 \cdot 10^{-13}$ , 0.446)
log(phonotactic probability)	-20.9	-19.9	11.3	(-61.6, -1.68)
Frequency	$2.82 \cdot 10^3$	$7.26 \cdot 10^2$	$7.73 \cdot 10^3$	(91, $1.13 \cdot 10^5$ )
log(frequency)	6.68	6.47	1.42	(2.25, 11.64)
Degree	18.7	15	14.7	(2, 75)
log(degree)	2.40	2.60	1.10	(-0.693, 4.30)
Clustering coefficient	0.34	0.31	0.18	(0, 1)
log(clustering coefficient)	-1.13	-1.07	0.642	(-3.40, 0.0328)
Coreness	11.13	9	9.17	(1, 47)
log(coreness)	2.03	2.14	0.878	(-0.693, 3.839)

<sup>7</sup> Note that “word length” in this table refers to the mean length in phones, computed over all pronunciations listed in the CMU dictionary. Thus, non-integer word lengths are possible. For example, the length of “are” is 1.5, since the CMU dictionary lists the pronunciations [ˈɑː] and [əː].

<sup>8</sup> To avoid taking  $\log(0)$ , a constant was added to degree, clustering coefficient, and coreness before applying a log transformation. For each variable the constant was equal to one half the difference between the lowest and second-lowest values of that variable. This step was taken for the other word properties so that all log transformations were carried out in the same way.

## Analysis

**Background: multilevel survival analysis.** At a high level, our goals in modeling this dataset are to understand how a set of variables affects when a given child produces a given word for the first time, and how the effects of some of these variables change over time. *Survival analysis* (also known as *event history analysis* or *hazard modeling*) is a statistical methodology commonly used to analyze this type of data: an event (such as a child producing the word *mommy*) is either observed to occur at, but not prior to time point  $t$ , or it is observed to never occur. A survival analysis models how the *hazard probability* of the event's first occurrence changes over time, according to a *hazard function*. The hazard function is also affected by a set of predictors (such as the degree of *mommy* in the child's input), whose effects may be allowed to change over time. Survival modeling is widely used in other fields to model events such as the onset of cigarette use (see e.g. Singer & Willett, 1991, 2003), and it has also been applied to language development datasets like ours, to model the first occurrence of various language milestones, e.g. the production of a child's first words (Ota & Green, 2013; Smolík, 2013; Tamis-LeMonda, Bornstein, & Baumwell, 2001; Tamis-LeMonda, Bornstein, Kahana-Kalman, Baumwell, & Cyphers, 1998).

Standard survival analysis assumes that observations are independent, conditional on the predictors. For structured data such as ours, which is grouped by child and by word, a multilevel version of survival modeling is needed, which takes into account the non-independence of observations from the same group, analogously to the multilevel regression models which are now widely used in language research (Baayen, Davidson, & Bates, 2008; Quené & van den Bergh, 2008). In particular, we use *multilevel discrete-time hazard modeling* (Barber et al., 2000; Reardon et al., 2002), which models the hazard of an event occurring at a discrete set of times, for grouped data. For our data, the event is a word being produced by a child for the first time in our sample. By including crossed random effects for word and child, we are able to model individual hazard

functions for each word and for each child. By including word-level and child-level variables in the model, we can assess how properties of words and of children affect the hazard of first production.

**Model specification.** Following Reardon, et al. (2002), the analysis was conducted as follows. First, a data matrix was prepared with one row per child per word per sampling age. For each child-word pairing, the dependent variable was set to one at the age at which the child first produced the word, and zero at time points prior to that age. Time points after first production were removed, because the event being modeled (first production) had already occurred. As noted above, if a child missed a recording session, all subsequent time points for that child (for all words) were removed. A discrete-time multilevel hazard model for this data can be estimated by fitting a multi-level logistic regression with a logit link (Barber et al., 2000). We did so using the `glmer()` function from the `lme4` package in R (Bates, Maechler, & Bolker, 2013), which fits multilevel generalized linear models using the Laplace approximation.

**Fixed effects.** A range of fixed effects were included in the model, to capture a number of factors which affect a child's probability of producing her first instance of a word at a given time. These are summarized in

Table 3. These predictors are related to several types of variables: (1) time (child's age during a session), (2) characteristics of children (child-level variables: gender, and caregiver lexical richness), and (3) properties of words (word-level variables). The word-level variables can be further divided into phonological network properties (degree, clustering coefficient, coreness), and other word properties (length in segments, frequency, phonotactic probability, part of speech). In addition to main-effect terms for variable types (1)-(3), terms for interactions between child age and word-level predictors were included, to examine changes in children's sensitivity to network structure over time.

**Table 3. Summary of predictors included in the analysis. "Type" describes the type of predictor ('cat': categorical; 'cont' continuous), along with the number of fixed-effect coefficients associated with it. Predictors which are**

residualized on others are indicated with a prime (e.g. DEGREE'). Centering, scaling (for continuous predictors) and contrast coding (for categorical predictors) are described in the text.

Term	Description	Type
<i>1. Time</i>		
AGE	Child age : second-order polynomial (two components)	cont(2)
<i>2. Child-level variables</i>		
FEMALE CHILD	Child is female	cat(1)
CAREGIVER LEXICAL RICHNESS	Mean word types per session produced by the caregiver in child-directed speech	cont(1)
<i>3. Word-level variables</i>		
<u>Network-theoretic properties</u>		
DEGREE'	Log of degree, residualized on log(length)	cont(1)
CLUSTERING COEFFICIENT	log of local clustering coefficient	cont(1)
CORENESS'	log of coreness, residualized on log(length) and log(degree)	cont(1)
<u>Other properties</u>		
LENGTH	log length in phonemes	cont(1)
FREQUENCY	log frequency in child-directed speech	cont(1)
PHON. PROBABILITY'	log of word's phonotactic probability, residualized on log(length)	cont(1)
PART OF SPEECH	part of speech (levels: noun, verb, modifier, closed class)	cat(4)
<i>4. Interactions</i>		
AGE:LENGTH	Interactions of word-level predictors with both components of AGE	cont(2)
AGE:DEGREE'		cont(2)
AGE:CLUSTERING COEFFICIENT		cont(2)
AGE:CORENESS'		cont(2)

**Child age** was coded in intervals of four months starting at 14 and ending at 50 months, and centered at 30 months. Based on visual inspection (e.g. Figure 2) and on fitting various baseline hazard models (including only a linear or non-linear function of age as a fixed effect), we determined that modeling the hazard's dependence on age using a restricted cubic spline with 3

knots provided the best fit (Harrell, 2001). The two spline components were then transformed into two principal components to reduce collinearity.

Among **word-level predictors**, the network-theoretic properties (degree, clustering coefficient, and coreness), as well as frequency and word length (in phones), were log-transformed to bring their distributions closer to normality (see Footnote 8). These variables are not all independent, raising the spectre that multicollinearity might mask the effects of network properties on when words are acquired, and lead to difficulty in fitting models. We thus took the following steps to minimize multicollinearity. Because log degree and log word length are highly correlated ( $r = -0.74, p < 0.001$  in our sample), we residualized log degree on log length. The resulting predictor, *DEGREE'*, tests whether degree has an impact on word learning *beyond* its shared variance with word length.<sup>9</sup> Likewise, log coreness is highly correlated with log degree ( $r = 0.94, p < 0.001$ ), as expected given that a word's coreness cannot, by definition, be greater than its degree. We therefore residualized log coreness on log degree, as well as log length (given the correlation between length and degree). The resulting predictor, *CORENESS'*, tests whether coreness has an impact on word learning beyond its shared variance with degree (and word length). log clustering coefficient (*CLUSTERING COEFFICIENT*) was not strongly related to the other three variables (Variance Inflation Factor = 1.5), so it was not residualized.

Since we were interested in changes in the impact of the phonological network properties over time, we included two-way **interactions** of *DEGREE'*, *CLUSTERING COEFFICIENT*, and *CORENESS'* with both of the child age components. Since two of these variables had been residualized on *LENGTH*, we also included the interactions of *LENGTH* with the age components.

In addition to the effects of network properties and their interactions with child age, which are of primary interest, we included several **word-level and child-level predictors**, to control for

---

<sup>9</sup> This residualization parallels the common practice in the neighborhood density literature, where separate analyses are conducted for words of each length (e.g. Charles-Luce & Luce, 1990, 1995; Coady & Aslin, 2003; but see Storkel, 2004, 2009).



known sources of variability in word learning. We included the child's gender and caregiver lexical richness (calculated as the mean number of word types produced per 90-minute recording session by the child's caregiver in child-directed speech, across all available samples), to account for advantages in language development for girls and for children who experience more child-directed speech (Hart & Risley, 1992; Hoff & Naigles, 2002; Hoff, 2003, 2010; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Rowe, 2008). We initially included SES as well, but its effect was not significant. As this measure was missing for six of the 62 children, we chose to drop it from the analysis, in order to include the data for those children.

At the word level, we included the word's log frequency in the present corpus of child-directed speech, log word length in phonemes, and log phonotactic probability, again relative to the present corpus. We expect a higher likelihood of short, frequent words being added to children's lexicons (e.g. Goodman et al., 2008; Storkel & Lee, 2011; Storkel, Maekawa, & Hoover, 2010; Storkel, 2004, 2009). We included phonotactic probability because of the possibility that observed advantages for high-density words in acquisition may in fact be due to their being composed of frequent phonological material (Coady & Aslin, 2003; Lindblom, 1992; Menn, 1978). Each word's phonotactic probability in the corpus of child-directed speech was calculated using a bigram model over phones, with bigram probabilities estimated from the caregiver speech corpus.<sup>10</sup> Because (log) phonotactic probability and (log) word length are highly correlated ( $r = -0.79$  in our sample), log phonotactic probability was residualized on log length. Before residualizing, log phonotactic probability was also significantly correlated with degree, clustering coefficient, and coreness ( $r = .64, .35, .69$ , respectively, all  $p < .001$ ), but after residualizing the correlations were small or non-significant ( $r = .12, p < .01, r = -.02, p > .7$ , and  $r = .04, p > .3$ , respectively). Finally, we also included

---

<sup>10</sup> Specifically: over the pronunciations as used in constructing the CDS phonological network, and including a word boundary symbol, unigram and and bigram probabilities were estimated using counts from the whole child-directed speech corpus, using Good-Turing smoothing to account for unseen bigrams. Each word's phonotactic probability was then calculated as the product, for each phone  $w_i$  in the word, of  $P(w_i | w_{i-1}) = P(w_i, w_{i-1})/P(w_{i-1})$  using the bigram and unigram models for  $P(w_i, w_{i-1})$  and  $P(w_{i-1})$ .

words' part of speech as a factor with four levels: noun, verb, modifier, and closed class, to account for a different baseline hazard of first production for each of these categories (Gentner & Boroditsky, 2001; Goodman et al., 2008).

To minimize collinearity, all continuous predictors were centered and all categorical predictors sum-coded. Continuous word-level and child-level predictors were centered at the word or child level (e.g. the mean log frequency across word *types* was subtracted from each word's log frequency). AGE was centered at 30 months. The condition number for the final set of predictors was 3.4, indicating minimal collinearity (Belsley, Kuh, & Welsch, 1980).

**Random effects.** The model included crossed random effects for words and for children. By-word and by-child random intercepts were included, to account for variability in children's "baseline" probability of producing new words and variability in words' baseline probability of being acquired, beyond the effects of child-level and word-level predictors included in the model. (The by-word random intercept can be more intuitively understood as capturing variability in when words tend to be acquired.)

The model also included all possible by-word and by-child random slopes corresponding to predictors of interest, given our modeling goals: AGE, network properties (DEGREE', CLUSTERING COEFFICIENT, and CORENESS'), their interactions with AGE, as well as LENGTH and its interaction with age (given that DEGREE' and CORENESS' were residualized on LENGTH). Thus, random slopes for AGE (both components) were included for both words and children. These random slopes, together with the random intercepts, capture variability among words and children in the baseline hazard function, including individual differences in children's rate of vocabulary growth as well as systematic variability across ages in children's talkativeness. Random slopes by child were included for DEGREE', CLUSTERING COEFFICIENT, CORENESS', AGE:DEGREE', AGE:CLUSTERING COEFFICIENT, AGE:CORENESS', LENGTH, and AGE:LENGTH, capturing variability among children in the influence of each of these word-level predictors on the baseline hazard function.

We included these random slopes to mitigate Type I error in the estimates of their fixed-effect coefficients (Barr, Levy, Scheepers, & Tily, 2013), but excluding correlations among the random effects.<sup>11</sup> Random slopes for the remaining predictors (child gender, caregiver lexical richness, frequency, phonotactic probability, and part of speech) were not included, because it proved unfeasible to fit models with all random slopes in a reasonable amount of time, and the effects of these predictors are not related to our primary modeling goals, so increased Type I error for their coefficients is acceptable. Accordingly, we will be tentative in drawing any conclusions from the fixed-effect coefficients for these predictors.

## Results

### Empirical Trends

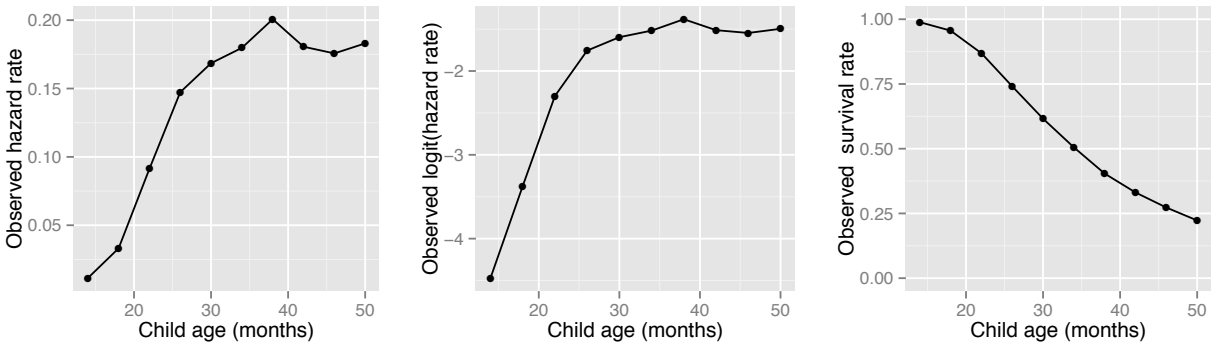
Before presenting the survival model of first productions just described, we examine plots of how children's likelihood of first producing a word, as well as the effect of network properties, change over time in the empirical data. These plots will give us a sense of what results to expect from the survival model.

Figure 2 shows the "overall" hazard function and survival function observed in our empirical data. The left panel plots the hazard at each time point  $t$ : the proportion of first production events which occur, out of all those which *could* occur, at a given age. This plot illustrates that the chance of a child producing a word for the first time rapidly rises for ages up to 30 months, after which it plateaus at 17.5-20%. The middle panel shows the logit (log-odds) of the hazard at each  $t$ , which is what is actually modeled in the survival model. The right panel shows the survival function: the proportion of word/child pairs for which acquisition has *not* occurred by time  $t$ , implied by the hazard function at times up to  $t$ . This survival function illustrates that the chance of

---

<sup>11</sup> Correlations between the random effects were excluded because the model failed to converge if correlations were included. Furthermore, because of the transformations already applied to the predictor set to minimize collinearity, the predictors of interest were largely uncorrelated, so we were confident that the variability in their effects among children and words would be as well.

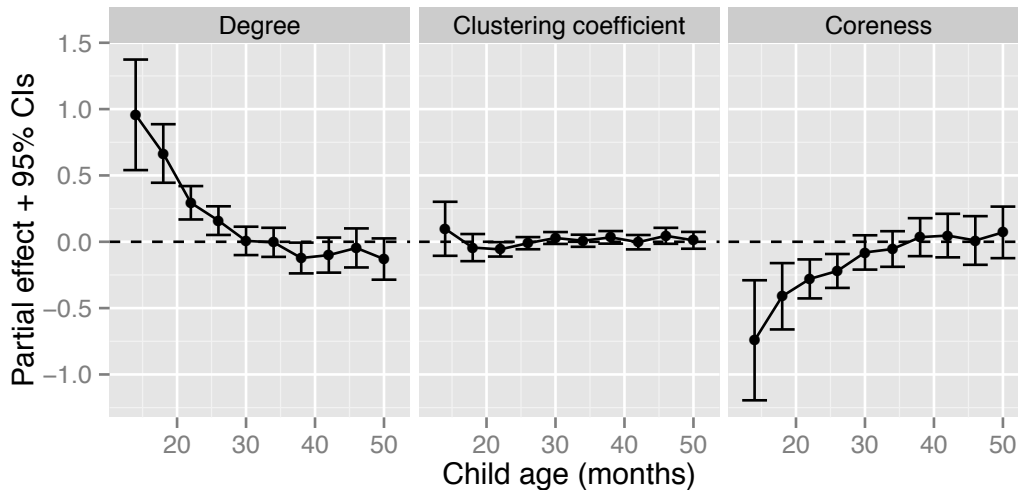
a word not having been acquired by a child, on average, decreases steadily over the study period, with the rate of decrease slowing slightly around 35 months.



**Figure 2.** Left: observed hazard of acquisition at each time point. Middle: same, with hazard expressed in log-odds (logit scale). Right: survival rate up to each time point, calculated using the observed hazard. (See text)

To visualize the empirical effects of the phonological network properties on the hazard of first production at each age, we used a more complex method to take into account collinearity between degree and coreness (and between both and word length). At each child age in the sample, we carried out a logistic regression of whether every possible first production event happened, using (log-transformed) degree, coreness, clustering coefficient, and word length (using the unresidualized versions of degree and coreness). The regression coefficient of each network property serves as a rough measure of how that property affects first production at a given age – without controlling for the many other factors (by-word and by-child offsets, word part of speech, etc.) included in the full survival model. Figure plots each network property’s regression coefficient, with 95% confidence intervals, over time. Degree (left panel) and coreness (right panel) start out having positive and negative effects, respectively, on the hazard of first production (higher-degree words and lower-coreness words are more likely to be produced). These effects rapidly diminish until about 30 months, after which degree and coreness have no effect. A word’s

clustering coefficient (middle panel) does not affect the likelihood of its being produced for the first time, at any age.



**Figure 3. Coefficients for degree, clustering coefficient, and coreness in logistic regressions of their joint effect (with word length) on the hazard of first production, carried out at each child age. (See text)**

### Survival model

We now turn to the results of the full survival model of the first production data, described above. Table 4 summarizes each fixed-effect coefficient corresponding to a predictor in the analysis (listed in

Table 3): its estimated value, standard error, and the corresponding Wald statistic and significance (by a Wald test). (The estimates of the random effect terms are listed in Appendix 1, but we do not discuss them further.) We discuss the model's predictions for how the hazard of a previously unuttered word being produced is affected by each type of predictor, in turn. The coefficients for AGE, which we consider first, determine the shape of the hazard function for an average child and word. The remaining fixed-effect coefficients can be interpreted as changing the position and shape of this hazard function: the main effects correspond to shifting the vertical position of the curve (more positive coefficient = acquisition at every time point is more likely), and interactions with age correspond to changing the shape of the curve.

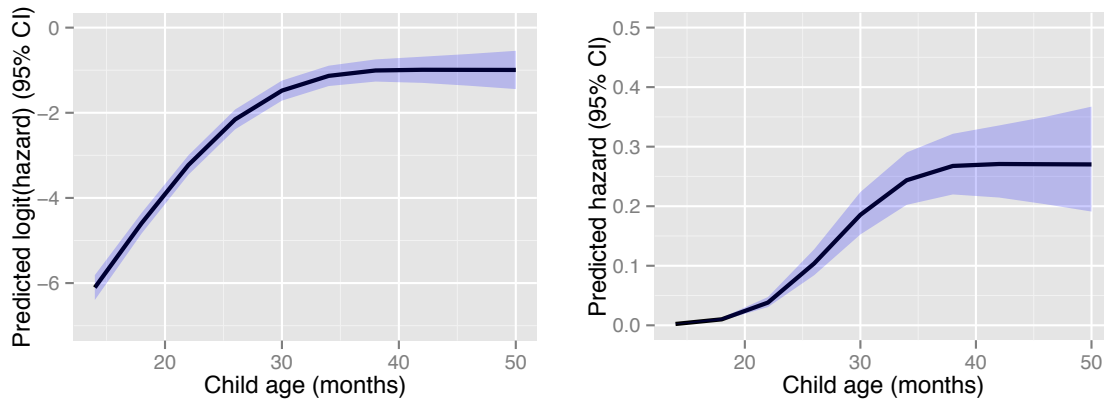
**Table 4. Summary of fixed effects for the model: coefficient estimates ( $\hat{\beta}$ ), standard errors, associated Wald z-scores, and significances (from a Wald test). Significances smaller than 0.01 are bolded. Predictors are grouped as in Table 3.**

Predictor	$\hat{\beta}$	SE( $\hat{\beta}$ )	z	p
Intercept	-2.96	0.11	-27.28	< <b>0.0001</b>
<i>1. Time</i>				
AGE <sub>1</sub> (component 1)	1.08	0.051	21.32	< <b>0.0001</b>
AGE <sub>2</sub> (component 2)	4.52	0.19	23.83	< <b>0.0001</b>
<i>2. Child-level variables</i>				
FEMALE CHILD	0.26	0.11	2.44	0.015
CAREGIVER VOCAB. SIZE	0.34	0.11	3.11	<b>0.0019</b>
<i>3. Word-level variables</i>				
<u>Network-theoretic properties</u>				
DEGREE'	0.22	0.051	4.25	< <b>0.0001</b>
CLUSTERING COEFFICIENT	-0.059	0.043	-1.36	0.17
CORENESS'	-0.38	0.14	-2.73	<b>0.0063</b>
<u>Other properties</u>				
LENGTH	-0.20	0.051	-3.88	<b>0.00011</b>
FREQUENCY	0.90	0.041	22.29	< <b>0.0001</b>
PHON. PROBABILITY'	0.0094	0.010	0.92	0.36
PART OF SPEECH = <i>noun</i> vs. <i>mean</i>	0.38	0.058	6.49	< <b>0.0001</b>
= <i>verb</i> vs. <i>mean</i>	-0.20	0.060	-3.29	<b>0.0010</b>
= <i>modifier</i> vs. <i>mean</i>	0.051	0.064	0.80	0.43
<i>4. Interactions</i>				
AGE <sub>1</sub> :LENGTH	-0.042	0.028	-1.50	0.13
AGE <sub>2</sub> :LENGTH	-0.097	0.10	-0.97	0.33
AGE <sub>1</sub> :DEGREE'	-0.15	0.029	-5.18	< <b>0.0001</b>

AGE <sub>2</sub> :DEGREE'	-0.46	0.11	-4.37	<b>&lt; 0.0001</b>
AGE <sub>1</sub> :CLUSTERING COEFFICIENT	0.054	0.024	2.20	0.028
AGE <sub>2</sub> :CLUSTERING COEFFICIENT	0.19	0.093	2.09	0.037
AGE <sub>1</sub> :CORENESS'	0.22	0.081	2.72	<b>0.0065</b>
AGE <sub>2</sub> :CORENESS'	0.34	0.29	1.18	0.24

**Age.** Because all variables have been centered (continuous) or sum-coded (categorical), the curve corresponding to the coefficients for AGE<sub>1</sub> and AGE<sub>2</sub> (plus the intercept) is the predicted hazard at each time, averaging across parts of speech and child genders, and holding the remaining variables at their mean values. Figure 4 shows this “overall hazard function”, in logit space (left) and probability space (right), with ribbons showing 95% confidence intervals.<sup>12</sup> This hazard function shows the likelihood that an average child will produce an average word at each age, *given that that child has not yet produced the word in our sample*. Thus, at age 14 months, the likelihood that a child will produce a previously unuttered word with average properties (in our sample) is predicted to start out near 0, then rapidly increase to around 25% by around 35 months, then stay at 25-28% until the end of sampling. This predicted overall hazard function can be compared to the trajectory of observed hazard in Figure 2. If the model fits the data well, these curves are expected to be similar – though not identical, since Figure 2 essentially shows an empirical mean which does not control for properties of words and children nor the grouping of observations by words and children, while the predicted hazard function comes from a model which does. The shapes of the observed and predicted curves in logit space are indeed very similar, with each one flattening out at 30-35 months. The predicted hazard function is somewhat “stretched” relative to the observed hazard, spanning a greater range in logit space (-6 to -1, versus -4.5 to -1.5). Nonetheless, the very good fit between the predicted and observed hazard functions gives some initial confidence in the predictions of our model.

<sup>12</sup> At each time point, the 95% CI was calculated using the variance-covariance matrix of the fixed-effect coefficient estimates.



**Figure 4.** Predicted overall hazard function, shown in logit space (left) and probability space (right). Solid line and shading are the predicted hazard (with all predictors besides child age held at their mean values), and its 95% confidence interval.

**Child-level variables, non-network theoretic word-level variables.** Consistent with previous research (Gentner & Boroditsky, 2001; Goodman et al., 2008; Hart & Risley, 1992; Hoff & Naigles, 2002; Huttenlocher et al., 1991; Rowe, 2008) we find robust advantages for girls and for children exposed to more lexically rich caregiver speech (FEMALE CHILD:  $\hat{\beta} = 0.26$ ,  $p = 0.015$ ; CAREGIVER LEXICAL RICHNESS:  $\hat{\beta} = 0.34$ ,  $p = 0.0019$ ). For example, the log-odds of a word's first production is increased by 0.52 for girls relative to boys ( $2 \times 0.26$ ), corresponding to a 68% increase in odds ( $\exp(0.52) = 1.68$ ), across all time points. We also found that shorter words and more frequent words are acquired earlier (Goodman et al., 2008; Storkel, 2004), with the hazard of first production higher for both types of words (LENGTH:  $\hat{\beta} = -0.20$ ,  $p = 0.0001$ ; FREQUENCY:  $\hat{\beta} = 0.90$ ,  $p < 0.0001$ ). As discussed above, the frequency effect controls for differences between words in the probability that children will need to use the word in the first place. We also find a significant difference between words with different parts of speech (PART OF SPEECH:  $F(3, 2101144) = 10.5$ ,  $p < 0.0001$ ). A post-hoc Tukey test showed that nouns had an advantage compared to modifiers ( $p = 0.003$ ), which had a marginal advantage over verbs ( $p = 0.067$ ), which did not differ significantly from closed-class words ( $p = 0.99$ ). Thus, the hazard of first production is highest for nouns, and



marginally higher for modifiers than for verbs and closed-class words, consistent with the noun bias found in previous research (e.g. Gentner & Boroditsky, 2001). However, unlike previous results (Storkel & Lee, 2011; Storkel, 2009), we did not detect a significant effect of phonotactic probability (after residualizing on length) (PHONOTACTIC PROBABILITY':  $p = 0.36$ ).

As noted above, the exact values of the estimates for these fixed effects, and their significances, may be inaccurate, due to the absence of associated random slopes. Nonetheless, the signs and significances of the estimated coefficients replicate established findings, capturing the associated variance in our data and supporting the validity of the survival analysis technique.

**Network-theoretic word-level variables.** We are primarily interested in the effects of phonological network properties on the hazard function, and how their effects change over time. These effects are captured in the model by three fixed-effect terms per network property (one main effect, two interactions with the two components of age), which together predict the “average effect” of the property (across time points, children, and words), and how the property affects the trajectory of the hazard function at each time point.

Beginning with the main effects, we find that the most local (degree) and the larger-scale (clustering coefficient and coreness) phonological network properties of words show opposite effects on word learning: high-degree words are more likely to be produced than low degree words (DEGREE':  $\hat{\beta}=0.22$ ,  $p < 0.0001$ ), and children are more likely to add *low* coreness words to their productive vocabularies than high coreness words (CORENESS:  $\hat{\beta}=-0.38$ ,  $p=0.0063$ ).<sup>13</sup> A lower local clustering coefficient is also associated with a word being produced for the first time, but the effect does not reach significance ( $p = 0.17$ ).

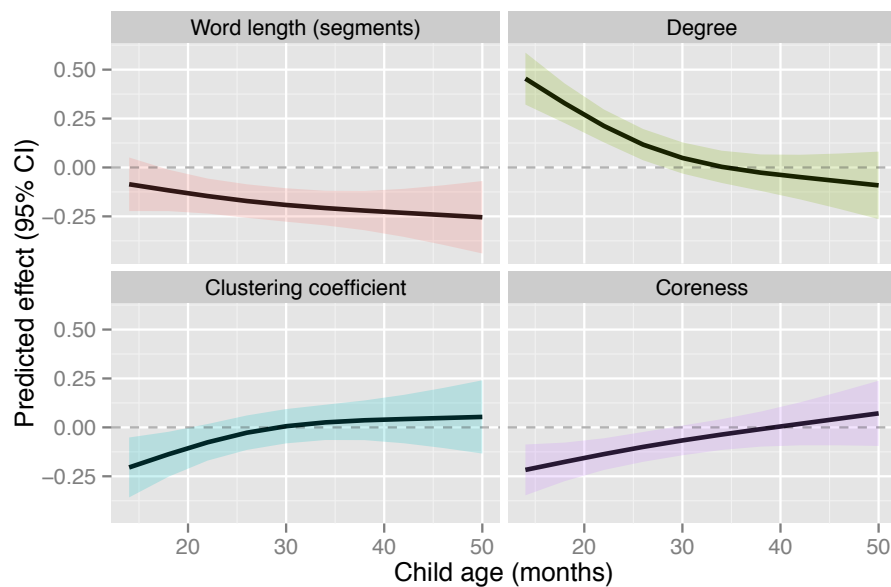
Turning to the interactions with AGE, we find that the effects of all three network properties on word learning change significantly over time (AGE<sub>1</sub>:DEGREE', AGE<sub>2</sub>:DEGREE' :  $p < 0.0001$ ;

AGE<sub>1</sub>:CLUSTERING COEFFICIENT :  $p = 0.028$ , AGE<sub>2</sub>:CLUSTERING COEFFICIENT :  $p = 0.037$ ; AGE<sub>1</sub>:CORENESS' :  $p <$

---

<sup>13</sup> Note that the effect of residualized coreness should be interpreted as a comparison of words with different coreness, but the same degree.

0.0065,  $AGE_2: CORENESS'$ :  $p = 0.24$ ). Figure 5 shows the development over time of the predicted effects of  $DEGREE'$ ,  $CLUSTERING COEFFICIENT$ , and  $CORENESS'$  on the hazard function – in addition to word length, on which degree and coreness were residualized – with all predictions standardized so that the y-axis corresponds to a change of 1 standard deviation in each network property. In other words, the heights of the curves in each panel can be interpreted as relative effect sizes. Positive y-values indicate that the predictor increases the likelihood of a word's entry into children's speech, and negative values indicate that the predictor decreases that likelihood.



**Figure 5. Predicted effect over time on log-odds of the hazard probability of each network property ( $DEGREE'$ ,  $CLUSTERING COEFFICIENT$ ,  $CORENESS'$ ), as well as  $LENGTH$ . (Note that  $DEGREE'$  and  $CORENESS'$  have been residualized on other variables, as described in Table 3.) The effects of all predictors have been standardized, so the y-axis corresponds to the predicted change in log-odds of hazard when the predictor is changed by one standard deviation. Solid lines and shading correspond to the predicted effect and its 95% confidence interval. Dotted lines indicate  $y=0$  (no effect), for clarity.**

Several patterns stand out from Figure 5. First, the effects of all three phonological network properties (degree, clustering coefficient, and coreness), are strongest at the earliest age observed,

and their strength diminishes over time, with all three effects vanishing by age 30 months (in the sense of the 95% CI of the effect intersecting with 0). Interestingly, this is also when the mean hazard curve (Figure 4) levels off, indicating a relatively uniform likelihood of children producing any particular new word for the first time after this age.

The effects of the network properties over time predicted by the model, shown in Figure 5, can be compared with the observed partial effect of each network property at each time point, shown in Figure 3. For degree and coreness, the observed and predicted patterns are very similar: a positive effect for degree and a smaller (in magnitude), negative effect for coreness, whose magnitudes are largest at the earliest age in the sample, and become negligible by around 30 months. For clustering coefficient, the observed pattern, of essentially no effect at all time points, differs from the model estimate of a small negative effect at early ages, which disappears by about 30 months. It is likely that the model is picking up on an effect of clustering coefficient which is only clear once word-level and child-level properties are controlled for (as is not done in the observed trajectories).

One might wonder if the attenuation of these effects with time as estimated by the model is inevitable, e.g. if children nearly exhaust the available words with certain properties, removing the possibility of observing a bias later on, even if the underlying effects remained the same. However, while this might restrict the range of certain properties in the remaining set of words to be acquired at later ages, this reduced range should only impact the standard errors of the estimates for the network properties, rather than the estimates themselves. Furthermore, the fact that the separate regressions at each time point in Figure 3 do not show increased standard errors over time suggests that reduced range for the network properties should not affect the model's results. Note also that the effect of word length does not attenuate with child age, suggesting that early preferences for words with certain properties do not necessarily lead to a reduction in the effects of those properties at later ages. Finally, refitting the model with all words that were produced at least

once by at least 10 children in the sample (compared to the cutoff of 30) revealed the same pattern of attenuation by 30 months, suggesting that this result is not an artifact of the reduced set of words selected here.

**Robustness.** Our model uses a relatively complex set of residualized predictors to parameterize network properties, for reasons discussed above. Given that the interpretation of results involving residualized predictors can be quite subtle, it is reasonable to wonder whether the model's core results – that three network properties *do* affect first productions, the directions of their effects, and how their effects change over time – are artifacts of using this particular set of residualized predictors, rather than showing that the three *unresidualized* network properties of interest (DEGREE, CORENESS, CLUSTERING COEFFICIENT) do affect first productions in the ways predicted by the model. We address this concern in two ways.

First, to provide a more stringent test that the three network properties each contributed independently to the model fit, we used likelihood ratio tests to compare a set of nested models, containing progressively more network properties:

- Model 1 : No network properties (DEGREE', CLUSTERING COEFFICIENT, CORENESS') or interactions with age (fixed- or random-effect terms).
- Model 2 : Model 1 + fixed and random-effect terms for DEGREE' and its interactions
- Model 3.1 : Model 2 + fixed and random-effect terms for CLUSTERING COEFFICIENT and its interactions
- Model 3.2 : Model 2 + fixed and random-effect terms for CORENESS' and its interactions
- Model 4: the full model presented above

Each model contained all other word and child-level predictors and interactions, plus the associated random-effects terms. These models can be nested either in the order  $1 < 2 < 3.1 < 4$ , or  $1 < 2 < 3.2 < 4$ . Either ordering takes degree to be preliminary in some sense to coreness and clustering coefficient: we first ask whether adding DEGREE' to a model with no network properties

improves the model, then whether CLUSTERING COEFFICIENT and CORENESS' improve the model on top of DEGREE', and finally whether adding both CLUSTERING COEFFICIENT and CORENESS' improve the model over the models with only one or the other. Note that because of how DEGREE' AND CORENESS' are defined, comparing Model 1 with Model 2 corresponds to asking whether *unresidualized* DEGREE contributes to model fit (beyond the effect of LENGTH), and comparing Model 2 with Model 3.1 similarly evaluates the contribution of unresidualized CORENESS.

Comparing Models 1 and 2 reveals that DEGREE' contributes significantly to the model fit, beyond all non-network predictors ( $\chi^2(6) = 46.99, p < .001$ ). Adding either CLUSTERING COEFFICIENT (Model 3.1;  $\chi^2(6) = 19.99, p < .01$ ) or CORENESS' (Model 3.2;  $\chi^2(6) = 24.30, p < .001$ ) to Model 2 also yielded significant improvements in model likelihood. Finally, the full model (Model 4) was found to be significantly superior to Model 3.1 ( $\chi^2(6) = 15.94, p < .05$ ), and marginally superior to Model 3.2 ( $\chi^2(6) = 11.63, p = .07$ ). The fact that adding CLUSTERING COEFFICIENT only had a marginal effect in this last comparison is consistent with its effect size and significance being weaker than that of DEGREE' and CORENESS', as seen in Table 4 and Figure 5, and with its absence in Figure 3. Overall, these likelihood ratio tests support the conclusion that degree and coreness, and to a lesser extent, clustering coefficient, have independent effects on the hazard of first production in our dataset.

Second, the good match discussed above at each time point between the model's predictions (cf. Figure 5), which use residualized network properties, and the observed effects of the network properties (cf. Figure 3), which do not, offers important reassurance that the direction of the effects of network properties predicted by our model, as well as how they change over time, are not artifacts of using residualized predictors.<sup>14</sup>

---

<sup>14</sup> We also attempted to fit a model using the unresidualized predictors, but the model failed to converge in a reasonable amount of time, probably due to the much more complex random effect structure necessitated by using unresidualized predictors.

### Discussion

We set out to explore larger-scale phonological network structure because of its implicit relevance to specific theoretical ideas concerning lexical and phonological development. Using longitudinal samples of child speech to estimate the age of first production for individual words, we examined two questions: whether and how local and larger-scale phonological network properties are related to children's productive vocabulary growth, and whether these effects change as a function of the child's age. Concerning the first question, we found that the timing of words' entry into children's productive lexicons is positively related to degree, the most local phonological network property of words, and negatively related to two measures of words' larger scale relationships with the rest of the lexicon, clustering coefficient and coreness. These results provide clear evidence that larger-scale network structure is related to vocabulary development. Concerning our second research question, all of these effects are present at 14 months, the earliest age sampled, and all have disappeared by age 30 months.

The facilitatory effect of degree reported here replicates the results of other studies on word learning under conditions resembling naturalistic contexts (Coady & Aslin, 2003; Demke et al., 2002; Hollich et al., 2002, Experiments 2 and 3; Merriman & Marazita, 1995; Storkel, 2004, 2009). Inhibitory effects are generally found when multiple novel, highly similar words are presented at the same time (Hollich et al., 2002, Experiment 1; Stager & Werker, 1997), but this situation may be unlikely in children's day-to-day experience. A positive degree effect is also consistent with the results of Storkel and Lee's (2011) delayed posttest, which they interpreted to indicate lexical support from phonological neighbors during configuration and engagement, the later stages of the development of phonological representations. Since our spontaneous production data are more likely to reflect the later stages of learning than children's early encounters with words, our findings offer support for Storkel and Lee's interpretation that degree positively affects these later stages by facilitating lexical processing, particularly working memory,

A positive effect of degree is also commensurable with the notion that sensitivity to phonological detail in lexical representations is driven by the need to distinguish among confusable words (Hallé & de Boysson-Bardies, 1996; Metsala & Walley, 1998; R. S. Newman, 2008; Vihman, 1996; Walley, 1993). While this idea might predict initial difficulty learning similar words (both in children's earliest productive vocabulary growth, and in their initial encounters with particular words), once children gain a foothold, the more rapid development of detailed representations (Garlock et al., 2001) in dense neighborhoods may make it easier to add new words to those neighborhoods than to sparser areas of the lexicon. That is, distinguishing a new target from many close neighbors may be easier if the child has easy access to the relevant details in her representations of the neighbors. Such a foothold may be gained by around 14 months, the age at which the present study commenced (Stager & Werker, 1997; Swingley & Aslin, 2002; Werker, Fennell, Corcoran, & Stager, 2002; Yoshida et al., 2009).

This interpretation of the facilitatory effect of degree makes intuitive sense, but does not fit well with the inhibitory effects of clustering coefficient and coreness. In fact, phonological detail in a new word's potential neighbors is not a function of the word's degree, but rather of the degree of these potential neighbors and the interconnectedness of the surrounding lexicon. Accelerated development of phonological detail in more interconnected areas would predict positive effects of clustering coefficient and coreness, not negative.

The present findings thus do not support an interpretation of phonological similarity effects based purely on the representational demands of distinguishing among similar words. They do, however, resonate with recent findings of inhibitory effects of clustering coefficient on lexical processing in adults (Altieri et al., 2010; Chan & Vitevitch, 2009, 2010; Yates, 2013), together with the well-established finding that degree facilitates word production (Vitevitch & Sommers, 2003; Vitevitch, 2002). These parallels between our present results and those of adult lexical processing

studies strongly encourage us to consider how children's productive vocabulary growth depends on processes associated with working memory and lexical access.

How might such an explanation work? One suggestion is to extend spreading-activation type models (Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994),<sup>15</sup> which account well for the facilitatory effect of degree on production, to account for the inhibitory effects of clustering coefficient (Chan & Vitevitch, 2009). When activation is allowed to flow not only between a target and its neighbors, but also between those neighbors and their neighbors, Chan and Vitevitch propose that high clustering coefficient may cause activation to pool among the neighbors because competitors pass activation to each other as well as to the target, increasing competition. When clustering coefficient is low, activation is either passed back to the target or dissipates to more distant words, decreasing competition. Network structure is thus one way in which the influence of phonological neighbors may be modulated by their relative activation compared to the target word (Chen & Mirman, 2012). An alternative interpretation is suggested by Gruenenfelder and Pisoni's (2009) observation that clustering coefficient reflects the extent to which a word differs from its neighbors at the same segmental position. Maximum clustering coefficient occurs when a word and its neighbors all differ at the same position, and minimum clustering coefficient occurs when a word's neighbors are most evenly spread across all segmental positions. Altieri, et al. (2010) used this observation to explain the inhibitory effects of clustering coefficient via feedback from the lexical to the segmental level.

Whether either of these accounts can explain the relationship between lexical processing and clustering coefficient is not yet clear, and while coreness can be expected to behave similarly to clustering coefficient—with activation pooling within more tightly knit subgroups within the network (Alvarez-Hamelin et al., 2005)—to date there has been no research into the psycholinguistic effects of coreness. The precise mechanisms underlying the role of larger-scale

---

<sup>15</sup> We see no reason why Shortlist-B (Norris & McQueen, 2008), modified to include larger-scale phonological network structure, should not make similar predictions.



phonological network structure in processing will thus require further study, but under this view, our present data from children's spontaneous speech do not so much reflect how children sample words for acquisition from their environment, or even how they manage to maintain independent long-term memory representations of words in their growing lexicon. Rather, we may interpret them as reflecting the process of learning to produce words with which the child already has ample receptive experience (Stokes, 2014). Children are more likely to produce words for the first time (increased hazard) when their position in the phonological network facilitates lexical processing for production (i.e. when degree is high, and clustering coefficient and, tentatively, coreness are low).

This raises the question of how to interpret the observed attenuation of these effects during the third year of life. Attributing the present findings to a processing-based account motivated by data from adults (Altieri et al., 2010; Chan & Vitevitch, 2009, 2010; Vitevitch & Sommers, 2003; Vitevitch, 2002; Yates, 2013) would seem to imply that phonological network effects on lexical production should not disappear, as we find here. While these effects can be detected experimentally in adults, they may nonetheless weaken with age, and this question should be explored further. But taking age as a rough proxy for vocabulary size suggests another possibility: it may simply be that these effects are no longer apparent in spontaneous speech data, once children have acquired a large enough lexicon. As noted above, word production for a child with a larger vocabulary likely depends more on what they need to say than on what they can easily say (Ota & Green, 2013; Rowland et al., 2008; Tomasello & Stahl, 2004). As a result, at later ages a longer time may elapse between the child first learning to produce a word and our ability to observe a first production in a spontaneous speech sample. The present results thus allow us to conclude that there is a practical attenuation of phonological network effects on words' first appearances in children's speech, but it is unclear whether the mechanism underlying these effects at earlier ages becomes less sensitive to phonological network structure at later ages. Ease of processing may yield measurable effects on children's spontaneous first productions while their lexicons are small, but

more controlled laboratory conditions may reveal continued effects in older children, as they do in adults.

The results of our inquiry thus make a compelling case that a crucial component of understanding how children acquire a lexicon needs to include a detailed account of how children's lexical processing develops. This must include both children's word recognition (Fernald, Swingley, & Pinto, 2001; Garlock et al., 2001; Metsala, Stavrinos, & Walley, 2008; Swingley & Aslin, 2000, 2007), as well as the processes underlying lexical production (e.g. Edwards, Beckman, & Munson, 2004; Munson, Swenson, & Manthei, 2005; R. S. Newman & German, 2002; Zamuner, Gerken, & Hammond, 2004). It must also, ultimately, be integrated with accounts of how children learn to articulate sounds and to associate words with meanings (Davis & MacNeilage, 1995; de Boysson-Bardies, Sagart, & Durand, 1984; Oller, Wieman, Doyle, & Ross, 1976; Özçaliskan & Goldin-Meadow, 2005).

Considering how children access words in their mental lexicon for comprehension and production, based on phonological network and other systematic lexical structure, will also require understanding the dynamic nature of that lexicon. Beyond the simple fact that children constantly learn new words, future research must also untangle how network structure changes in response to the gradual nature of word learning. Progressive elaboration of phonological detail in lexical representations (Metsala & Walley, 1998; Walley, 1993), the movement of words from receptive to productive vocabulary (Stokes, 2014), and progress through stages such as triggering, configuration, and engagement (Storkel & Lee, 2011), may all impact how words participate in the phonological network. This points out a need to evaluate carefully how we select a lexicon in which to calculate phonological network properties. We have used the pooled caregiver speech samples in the Chicago corpus, which offers a reasonable approximation of the lexical input available to English-speaking children in Chicago. However, the processing-based interpretation of our results requires measuring phonological network properties in children's own lexicons.

A practical response to this need is to consider child-directed speech as an approximation of the child's receptive lexicon. This is neither a new idea (e.g. Coady & Aslin, 2003) nor an unreasonable one, because by 14 months children have accumulated many months of experiencing child-directed speech. While they may not yet reliably comprehend most words, they are likely to be quite familiar with many, and well on their way to becoming familiar enough with others. We have attempted to capture some of the variability in children's familiarity with specific words by controlling for frequency in our survival model.

Another approach might use children's prior productions to define a productive lexicon at each time point. This would enable us to explore whether children are more likely to produce new words based on their connectivity with other words *that the child already produces* (e.g. the analysis of semantic network structure in Hills et al., 2009). However, it is not clear that the set of "words known to a child" at a given point in time should be restricted to those words that the child knows how to produce. This restriction could help determine if the ability to articulate certain words helps children learn to produce similar words, but this is only one facet of the role that phonological representations play in lexical development. The theoretical contributions reviewed above (Metsala & Walley, 1998; Stokes, 2014; Storkel & Lee, 2011; Walley, 1993) point to a more nuanced characterization of children's lexicons taking into account children's varying familiarity with words, from previously unencountered words, to words that children recognize but do not yet produce, to words that children produce reliably. We leave this important issue to future work.

One further issue that bears mention is morphology. In the interest of modeling the complete lexicon of phonological forms we have included inflections as independent nodes in the network. Morphological relationships among words, of course, may also support word learning, but in many cases English inflections also differ by only one phoneme (cf. *walk, walks, walked*), potentially confounding the influence of phonological and morphological relationships on first productions. One strategy for separating these variables has been to exclude inflections, at least

after the child has acquired the bare root (e.g. Coady & Aslin, 2003), but taking this approach would necessitate several decisions about the contents of phonological neighborhoods, e.g. whether to include only neighbors of a single root form, or to include all neighbors of any member of the inflectional paradigm. However, inflectional paradigms are small in English, and single-phoneme inflections (primarily plurals and tense markers) affect mainly nouns and verbs. The inclusion of inflections in the network thus affects most nouns and verbs in a similar way, such that any confounding of morphological and phonological network structure is likely to be captured by the inclusion of lexical category in our analysis.

We close our discussion with a note about individual differences among children. The survival model revealed systematic effects of the phonological network properties across children (Table 4, Figure 5), but the estimates for the random-effects terms (Appendix 1) suggest meaningful differences between children in their sensitivity to phonological network structure (see also Maekawa & Storkel, 2006; Stokes, 2014; and similar reasoning related to semantic networks in Beckage, Smith, & Hills, 2011), as well as substantial additional variance in children's vocabulary growth that we have not accounted for. Apart from including as covariates gender and amount of caregiver input, we leave investigating these individual differences for future research.

### **Conclusion**

Our survival analysis of words' first appearances in the spontaneous speech of children aged 14-50 months has allowed us to investigate the role of local and larger-scale phonological network structure in vocabulary acquisition. We have demonstrated that a word's degree, corresponding to the traditional concept of neighborhood density, is positively related to the likelihood of observing a previously unuttered word in children's speech, corroborating prior results (Coady & Aslin, 2003; Storkel, 2009). We have also shown that two larger-scale network properties that had not previously been examined in child lexical development, clustering coefficient and coreness, are negatively related to children's likelihood of uttering a previously

unobserved word, and that the effects of all three phonological network properties attenuate by age 30 months, at least in spontaneous speech data. By controlling for a variety of other factors and carefully addressing interrelationships among the various lexical properties (e.g. residualizing on word length), we have shown that these network properties contribute to the hazard function above and beyond the other influences examined.

This pattern of results strongly resembles the effects of degree and clustering coefficient (as well as coreness, which, we argue, can be expected to pattern with clustering coefficient) on adult speech production, suggesting that learning to produce words is dependent on processes that underlie lexical access for speech production (Storkel & Lee, 2011). Therefore, exploring how models of adult lexical processing may be adapted to account for child data, and understanding how these processes develop, coupled with a robust understanding of how children deploy these processes under different task conditions (Werker & Curtin, 2005) will be crucial to advancing our understanding of lexical development.

Finally, the application of network science to psycholinguistics and language acquisition is still very young (excepting the long tradition of research on one small aspect of network structure, neighborhood density). By placing the focus on how network structure modulates children's access to words for production, the present results are a strong reminder that our primary goal is not to understand the structural organization of the mental lexicon—it is to understand how children come to know *and use* a lexicon (cf. Stoel-Gammon, 2011). Our results show how phonological network structure can help us understand the interplay between phonological form and this process.

**Acknowledgments**

Data analysis and preparation of this article were completed with the support of a NAEd/Spencer Foundation postdoctoral fellowship to MTC, and by a Daniels Graduate Fellowship at the University of Chicago to MS. We express our sincerest gratitude to Susan Goldin-Meadow for making the Chicago Corpus, collected under NICHD grant P01 HD40605 to SGM, available to us. We also wish to thank Stephen Raudenbush and Jacob Foster for suggestions regarding the analysis; Jason Voigt and Kristi Schonwald for their assistance with the corpus; Matt Goldrick, Melinda Fricke, Mits Ota, and two anonymous *JML* reviewers for their helpful comments on the manuscript; as well as audiences at the University of Chicago, Northwestern University, NetSci 2012, and the LSA for their feedback.

## References

- Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, 74, 47–97.
- Altieri, N., Gruenenfelder, T., & Pisoni, D. B. (2010). Clustering coefficients of lexical neighborhoods: Does neighborhood structure matter in spoken word recognition? *The Mental Lexicon*, 5(1), 1–21.  
doi:10.1075/ml.5.1.01alt
- Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). Large scale networks fingerprinting and visualization using the k-core decomposition. In *Advances in neural information processing systems* (pp. 41–50). Retrieved from [http://cnet.fi.uba.ar/ignacio.alvarez-hamelin/pdf/AH\\_D\\_B\\_V\\_NIPS2006.pdf](http://cnet.fi.uba.ar/ignacio.alvarez-hamelin/pdf/AH_D_B_V_NIPS2006.pdf)
- Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2008). k-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases. *Networks and Heterogeneous Media*, 3(2), 371–393. doi:10.3934/nhm.2008.3.371
- Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), 11149–11152. doi:VL - 97
- Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010a). Comparative analysis of networks of phonologically similar words in English and Spanish. *Entropy*, 12(3), 327–337. doi:10.3390/e12030327
- Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010b). The structure of phonological networks across multiple languages. *International Journal of Bifurcations and Chaos*, 20(3), 679–685.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Bailey, T. M., & Hahn, U. (2005). Phoneme similarity and confusability. *Journal of Memory and Language*, 52(3), 347–370.
- Barber, J. S., Murphy, S. A., Axinn, W. G., & Maples, J. (2000). Discrete-Time Multilevel Hazard Analysis. *Sociological Methodology*, 30(1), 201–235. doi:10.1111/0081-1750.00079
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.

- Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using Eigen and syntax (R package Version 0.999999-2).
- Beckage, N., Smith, L., & Hills, T. (2011). Small Worlds and Semantic Network Growth in Typical and Late Talkers. *PloS One*, 6(5), e19348.
- Beckman, M. E., Munson, B., & Edwards, J. (2007). Vocabulary growth and developmental expansion of types of phonological knowledge. *Laboratory Phonology*, 9, 241–264.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.
- Bernstein Ratner, N., Newman, R., & Streckas, A. (2009). Effects of word frequency and phonological neighborhood characteristics on confrontation naming in children who stutter and normally fluent peers. *Journal of Fluency Disorders*, 34(4), 225–241. doi:10.1016/j.jfludis.2009.09.005
- Carlson, M. T., Bane, M., & Sonderegger, M. (2011). Global properties of the phonological networks of child and child-directed speech. In N. Danis, K. Mesh, & H. Sung (Eds.), *Proceedings of the 35th Annual Boston University Conference on Language Development* (pp. 97–109). Somerville, MA: Cascadilla Press.
- Carnegie Mellon Speech Group. (1993). The Carnegie Mellon Pronouncing Dictionary (Version cmudict.0.7a). Carnegie Mellon University.
- Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering-coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 35(6), 1934–1949.
- Chan, K. Y., & Vitevitch, M. S. (2010). Network structure influences speech production. *Cognitive Science*, 34(4), 685–697. doi:10.1111/j.1551-6709.2010.01100.x
- Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, 17(1), 205–215.
- Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language*, 22(3), 727–735.



- Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*(2), 417–430. doi:10.1037/a0027175
- Coady, J. A., & Aslin, R. N. (2003). Phonological neighborhoods in the developing lexicon. *Journal of Child Language*, *30*(2), 441–469. doi:10.1017/S0305000903005579
- Davis, B. L., & MacNeilage, P. F. (1995). The Articulatory Basis of Babbling. *Journal of Speech and Hearing Research*, *38*(6), 1199–1211.
- De Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, *11*(1), 1–15.  
doi:10.1017/S0305000900005559
- Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production* (pp. 9–37). Berlin: Mouton de Gruyter.
- Demke, T. L., Graham, S. A., & Siakaluk, P. D. (2002). The influence of exposure to phonological neighbours on preschoolers' novel word production. *Journal of Child Language*, *29*(02), 379–392.
- Dollaghan, C. A. (1994). Children's phonological neighbourhoods: half empty or half full? *Journal of Child Language*, *21*(2), 257–271.
- Dorogovtsev, S. N., Goltsev, A. V., & Mendes, J. F. F. (2006). K-core organization of complex networks. *Physical Review Letters*, *96*(4), 40601.
- Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*(2), 421.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*(2-3), 245–264.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5).
- Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*(2), 419–439.

- Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*(4), 1003–1015.
- Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45*(3), 468–492.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge: Cambridge University Press.
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard Words. *Language Learning and Development, 1*(1), 23–64. doi:10.1207/s15473341lld0101\_4
- Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35*(3), 515–531. doi:10.1017/S0305000907008641
- Gruenenfelder, T., & Pisoni, D. B. (2009). The lexical restructuring hypothesis and graph theoretic analyses of networks based on random lexicons. *Journal of Speech, Language and Hearing Research, 52*(3), 596–609.
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference (SciPy2008)* (pp. 11–15). Pasadena, CA USA.
- Hallé, P. A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development, 19*(4), 463–481.
- Harrell, F. E. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer.
- Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology, 28*(6), 1096–1096.
- Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language, 63*(3), 259–273.

- 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. doi:10.1111/j.1467-9280.2009.02365.x
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378.
- Hoff, E. (2010). Context effects on young children's language use: The influence of conversational setting and partner. *First Language, 30*(3-4), 461–472. doi:10.1177/0142723710370525
- Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development, 73*(2), 418–433.
- Hollich, G., Jusczyk, P. W., & Luce, P. A. (2002). Lexical neighborhood effects in 17-month-old word learning. In *Proceedings of the 26th annual Boston University conference on language development* (pp. 314–323). Boston: Cascadilla.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*(2), 236–248.
- Kleinberg, J. M. (2000). Navigation in a small world. *Nature, 406*, 845.
- Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure or equivalence assumptions for theories of word recognition. *Journal of Verbal Learning & Verbal Behavior, 12*, 119–131.
- Lindblom, B. (1992). Phonological units as adaptive emergents of lexical development. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications*. Timonium, MD: York Press.
- Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes, 16*(5/6), 565–581.
- Luce, P. A., & Pisoni, D. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear & Hearing February 1998, 19*(1), 1–36.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (Vol. 3rd). Mahwah: Lawrence Erlbaum.
- Maekawa, J., & Storkel, H. L. (2006). Individual differences in the influence of phonological characteristics on expressive vocabulary development by young children. *Journal of Child Language, 33*(3), 439–459.

- Mathey, S., Robert, C., & Zagar, D. (2004). Neighbourhood distribution interacts with orthographic priming in the lexical decision task. *Language and Cognitive Processes, 19*(4), 533–560.  
doi:10.1080/01690960344000260
- Mathey, S., & Zagar, D. (2000). The neighborhood distribution effect in visual word recognition: Words with single and twin neighbors. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 184–205. doi:10.1037/0096-1523.26.1.184
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1–86.
- McKean, C., Letts, C., & Howard, D. (2013). Functional reorganization in the developing lexicon: separable and changing influences of lexical and phonological variables on children's fast-mapping. *Journal of Child Language, 40*(2), 307–335.
- Menn, L. (1978). Phonological units in beginning speech. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments*. Amsterdam: North Holland.
- Merriman, W. E., & Marazita, J. M. (1995). The effect of hearing similar-sounding words on young 2-year-olds' disambiguation of novel noun reference. *Developmental Psychology, 31*(6), 973–984.
- Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition, 25*(1), 47–56.
- Metsala, J. L., Stavrinos, D., & Walley, A. C. (2008). Children's spoken word recognition and contributions to phonological awareness and nonword repetition: A 1-year follow-up. *Applied Psycholinguistics, 30*(1), 101–121.
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum Associates.
- Munson, B., Swenson, C. L., & Manthei, S. C. (2005). Lexical and phonological organization in children: evidence from repetition tasks. *Journal of Speech, Language, and Hearing Research, 48*(1), 108–124.
- Newman, M. E. (2003). The structure and function of complex networks. *SIAM Review, 45*(2), 167–256.

- Newman, R. S. (2008). The level of detail in infants' word learning. *Current Directions in Psychological Science*, *17*(3), 229–232.
- Newman, R. S., & German, D. J. (2002). Effects of Lexical Factors on Lexical Access among Typical Language-Learning Children and Children with Word-Finding Difficulties. *Language and Speech*, *45*(3), 285–317. doi:10.1177/00238309020450030401
- Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234. doi:10.1016/0010-0277(94)90043-4
- Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395. doi:10.1037/0033-295X.115.2.357
- Oller, D. K., Wieman, L. A., Doyle, W. J., & Ross, C. (1976). Infant babbling and speech. *Journal of Child Language*, *3*(1), 1–11. doi:10.1017/S0305000900001276
- Ota, M., & Green, S. J. (2013). Input frequency and lexical variability in phonological development: a survival analysis of word-initial cluster production. *Journal of Child Language*, *40*(3), 539–566. doi:10.1017/S0305000912000074
- Özçaliskan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, *96*(3), B101–B113.
- Pater, J., Stager, C., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*(3), 384–402.
- Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*(4), 413–425.
- Reardon, S. F., Brennan, R. T., & Buka, S. L. (2002). Estimating Multi-Level Discrete-Time Hazard Models Using Cross-Sectional Data: Neighborhood Effects on the Onset of Adolescent Cigarette Use. *Multivariate Behavioral Research*, *37*(3), 297. doi:10.1207/S15327906MBR3703\_1
- Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(1). doi:10.1017/S0305000907008343
- Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in Early Gesture Explain SES Disparities in Child Vocabulary Size at School Entry. *Science*, *323*, 951–953. doi:10.1126/science.1167025

- Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The Pace of Vocabulary Growth Helps Predict Later Vocabulary Skill. *Child Development, 83*(2), 508–525. doi:10.1111/j.1467-8624.2011.01710.x
- Rowland, C., Fletcher, S. L., & Freudenthal, D. (2008). How big is enough? Assessing the reliability of data from naturalistic samples. In *Trends in corpus research: Finding structure in data* (pp. 1–24). Amsterdam: John Benjamins.
- Sagae, K., Davis, E., Lavie, A., MacWhinney, B., & Wintner, S. (2010). Morphosyntactic Annotation of CHILDES Transcripts. *Journal of Child Language, 37*(Special Issue 03), 705–729. doi:10.1017/S0305000909990407
- Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110*(2), 268.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.
- Smolík, F. (2013). Noun Imageability Facilitates the Acquisition of Plurals: Survival Analysis of Plural Emergence in Children. *Journal of Psycholinguistic Research, 1*–16. doi:10.1007/s10936-013-9255-5
- Stager, C. L., & Werker, J. L. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381.
- Stevens, J., Yang, C., Trueswell, J., & Gleitman, L. (2012, January 5). *A Psychologically Motivated Model of Word Learning*. Paper presented at the Workshop on Psycho-computational Models of Language Acquisition, Portland, Oregon.
- 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.
- Stoel-Gammon, C. (2011). Relationships Between Lexical and Phonological Development in Young Children. *Journal of Child Language, 38*(1), 1–34. doi:10.1017/S0305000910000425
- Stokes, S. F. (2010). Neighborhood Density and Word Frequency Predict Vocabulary Size in Toddlers. *Journal of Speech Language, and Hearing Research, 53*(3), 670–683. doi:10.1044/1092-4388(2009/08-0254)
- Stokes, S. F. (2014). The impact of phonological neighborhood density on typical and atypical emerging lexicons. *Journal of Child Language, 41*(03), 634–657. doi:10.1017/S030500091300010X

- Stokes, S. F., Kern, S., & Dos Santos, C. (2012). Extended Statistical Learning as an account for slow vocabulary growth. *Journal of Child Language*, *39*(1), 105–130.
- Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, *25*(2), 201–221.
- Storkel, H. L. (2009). Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *Journal of Child Language*, *36*(2), 291–321.  
doi:10.1017/S030500090800891X
- 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*, *49*(6), 1175.
- Storkel, H. L., & Lee, S.-Y. (2011). The independent effects of phonotactic probability and neighbourhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, *26*(2), 191–211. doi:10.1080/01690961003787609
- Storkel, H. L., Maekawa, J., & Hoover, J. R. (2010). Differentiating the Effects of Phonotactic Probability and Neighborhood Density on Vocabulary Comprehension and Production: A Comparison of Preschool Children With Versus Without Phonological Delays. *J Speech Lang Hear Res*, *53*(4), 933–949.  
doi:10.1044/1092-4388(2009/09-0075)
- Swingle, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166.
- Swingle, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480–484.
- Swingle, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*(2), 99–132. doi:10.1016/j.cogpsych.2006.05.001
- Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal Responsiveness and Children's Achievement of Language Milestones. *Child Development*, *72*(3), 748–767. doi:10.1111/1467-8624.00313
- Tamis-LeMonda, C. S., Bornstein, M. H., Kahana-Kalman, R., Baumwell, L., & Cyphers, L. (1998). Predicting variation in the timing of language milestones in the second year: an events history approach. *Journal of Child Language*, *25*(3), 675–700.

- Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: how much is enough? *Journal of Child Language*, *31*(1), 101–121. doi:10.1017/S0305000903005944
- Vihman, M. M. (1996). *Phonological Development: The origins of language in the child*. Oxford: Blackwell.
- Vihman, M. M., & Velleman, S. L. (2000). The construction of a first phonology. *Phonetica*, *57*, 255–266.
- Vitevitch, M. S. (2002). The Influence of Phonological Similarity Neighborhoods on Speech Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 735–747.
- Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language and Hearing Research*, *51*(2), 408–442.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325–329.
- 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.
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1), 306–311.
- Vitevitch, M. S., & Sommers, M. S. (2003). The Facilitative Influence of Phonological Similarity and Neighborhood Frequency in Speech Production in Younger and Older Adults. *Memory & Cognition*, *31*(4), 491–504.
- Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, *13*, 286–350.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*, 440–442. doi:10.1038/30918
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234.
- Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1–30.
- Yarkoni, T., Balota, D., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979. doi:10.3758/PBR.15.5.971



Yates, M. (2013). How the Clustering of Phonological Neighbors Affects Visual Word Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(5), 1649–1656.

doi:10.1037/a0032422

Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, 12(3), 412–418.

Zamuner, T. S. (2009). The structure and nature of phonological neighbourhoods in children's early lexicons.

*Journal of Child Language*, 36(1), 3–21. doi:10.1017/S0305000908008829

Zamuner, T. S., Gerken, L., & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language*, 31(3), 515–536. doi:10.1017/S0305000904006233

**Appendix 1: Random effects**

**Table A1. Estimated variance ( $\hat{\sigma}^2$ ) and standard deviation corresponding to each by-child and by-word random-effect term included in the fitted model. For example,  $\hat{\sigma}^2=0.62$  for the by-child random intercept, and  $\hat{\sigma}^2=0.0030$  for the by-child random slope for DEGREE'.**

Group	Variable	Est. variance	Est. SD
Child	INTERCEPT	0.62	0.79
	AGE <sub>1</sub>	0.12	0.35
	AGE <sub>2</sub>	1.7	1.3
	LENGTH	0.0062	0.079
	DEGREE'	0.0030	0.055
	CLUSTERING COEFFICIENT	0.00034	0.019
	CORENESS'	0.0145	0.12
	AGE <sub>1</sub> :LENGTH	0.00070	0.026
	AGE <sub>2</sub> :LENGTH	0.0027	0.052
	AGE <sub>1</sub> :DEGREE'	6.8e-12	2.6e-06
	AGE <sub>2</sub> :DEGREE'	0.016	0.012
	AGE <sub>1</sub> :CLUSTERING COEFFICIENT	0	0
	AGE <sub>2</sub> :CLUSTERING COEFFICIENT	0.028	0.017
	AGE <sub>1</sub> :CORENESS'	0.00057	0.024
AGE <sub>2</sub> :CORENESS'	2.9e-10	0.000017	
Word	INTERCEPT	0.53	0.73
	AGE <sub>1</sub>	0.17	0.41
	AGE <sub>2</sub>	1.15	1.1

**Appendix 2: Word list**

again	bee	break	careful	crayons	duck
all	beep	bring	carry	crying	ear
already	being	broke	cars	cup	ears
am	better	brown	cat	cut	eat
an	big	bubble	catch	d	eating
and	bigger	bubbles	chair	dad	egg
any	bike	bug	check	daddy	eight
are	bird	build	cheese	dance	else
arm	bit	bump	chicken	dark	end
around	bite	bunny	circle	day	every
at	black	bus	clean	did	eyes
ate	block	but	climb	different	face
away	blocks	butt	close	dirty	fall
b	blow	button	cold	do	falling
babies	blue	buy	color	does	fast
baby	boat	by	colors	dog	feel
back	boo	c	come	doggie	feet
bad	book	cake	comes	dogs	fell
bag	books	call	coming	doing	find
ball	boom	called	cookie	done	fine
balloon	boots	came	cookies	door	fire
balls	both	camera	cool	dora	first
bath	bottle	can	couch	down	fish
be	bowl	candy	could	draw	fit
bear	box	cant	count	drink	five
because	boy	car	cow	drive	fix
bed	bread	card	crayon	drop	floor

flower	green	his	kids	lot	moon
fly	guess	hit	kind	lots	more
flying	guy	hold	kiss	loud	mouse
food	guys	hole	kitchen	love	mouth
foot	h	home	kitty	lunch	move
for	had	horse	knock	ma	movie
found	hair	hot	know	mad	much
four	hand	house	last	made	my
friend	hands	how	later	make	name
friends	happened	hug	lay	makes	need
from	happy	hurt	leave	making	needs
fun	hard	hurts	left	mama	never
funny	has	ice	leg	man	new
game	hat	if	legs	many	next
games	have	ill	let	maybe	nice
gave	having	in	lets	me	night
get	he	into	letter	mean	nine
gets	head	is	light	mess	noise
getting	hear	it	like	middle	nose
girl	heart	its	likes	might	not
give	heavy	job	lion	milk	now
go	help	juice	little	mine	number
goes	her	jump	live	miss	of
going	here	jumping	long	mom	off
gone	hide	just	look	mommy	old
good	hiding	keep	looking	money	on
got	high	key	looks	monkey	one
grandma	him	kick	lost	monster	ones

only	play	sad	sitting	stuff	these
open	playing	said	six	sun	they
or	pop	same	sleep	supposed	thing
other	potty	sandwich	sleeping	sure	things
our	press	save	slide	swing	think
out	pretty	saw	small	t	this
over	pull	say	snack	table	those
own	puppy	says	so	tail	though
page	purple	scared	sock	take	thought
paint	push	scary	some	taking	three
pants	put	school	song	talk	through
papa	putting	seat	sorry	talking	throw
paper	puzzle	see	spider	tape	tickle
park	rain	set	spoon	taste	tiger
part	reach	seven	square	teeth	time
party	read	she	stand	tell	tired
people	ready	shirt	star	ten	to
phone	real	shoe	start	than	today
pick	really	shoes	stay	thank	toe
picture	red	should	step	that	told
pictures	remember	show	stick	the	tomorrow
piece	ride	side	sticker	their	too
pieces	right	silly	still	them	took
pig	rock	sing	stop	then	top
pillow	roll	sit	stuck	there	touch
pink	room				
place	run				
plate					

towel	wants	with
tower	was	wont
toy	wash	work
toys	watch	working
train	watching	works
tree	water	would
truck	way	write
try	we	wrong
trying	wear	yellow
tummy	went	yet
turn	were	you
turned	wet	your
turtle	what	yours
two	whats	yucky
um	when	yum
under	where	yummy
up	which	
us	white	
use	who	
very	whole	
wait	why	
wake	will	
walk	win	
walking	window	
want	wipe	