



Semantic facilitation in bilingual first language acquisition [☆]



Samuel Bilson ^a, Hanako Yoshida ^b, Crystal D. Tran ^b, Elizabeth A. Woods ^b, Thomas T. Hills ^{c,*}

^a *Institute of Mathematics, University of Warwick, United Kingdom*

^b *Department of Psychology, University of Houston, United States*

^c *Department of Psychology, University of Warwick, United Kingdom*

ARTICLE INFO

Article history:

Received 31 August 2014

Revised 18 February 2015

Accepted 29 March 2015

Keywords:

Bilingualism

Language acquisition

Communicative development inventory

Network analysis

Mutual exclusivity

ABSTRACT

Bilingual first language learners face unique challenges that may influence the rate and order of early word learning relative to monolinguals. A comparison of the productive vocabularies of 435 children between the ages of 6 months and 7 years—181 of which were bilingual English learners—found that monolinguals learned both English words and all-language concepts faster than bilinguals. However, bilinguals showed an enhancement of an effect previously found in monolinguals—the preference for learning words with more associative cues. Though both monolinguals and bilinguals were best fit by a similar model of word learning, semantic network structure and growth indicated that the two groups were learning English words in a different order. Further, in comparison with a model of two-monolinguals-in-one-mind, bilinguals overproduced translational equivalents. Our results support an emergent account of bilingual first language acquisition, where learning a word in one language facilitates its acquisition in a second language.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

How does learning two first languages at the same time differ from learning only one first language? This question is the focus of this article but also holds a central place in our understanding of the cognitive prerequisites for language learning more generally. From a casual perspective, learning two languages at once should make the language learning problem harder, because words and their referents violate the one-to-one mapping we often expect within a language. Most bilingual children, however, seem undaunted. The evidence suggests that bilingual first language learners have little problem learning two languages

(Lanza, 2000), and as we report below, past research suggests that the differences are, if anything, subtle. Yet, these differences are important because our theories of language acquisition rely on certain underlying principles, and many of these principles are derived from research on monolinguals. By understanding bilingual learning better, we better understand the constraints and generality of our theories. We also better understand what is both a growing form of language learning (UNESCO, 2003) as well as a language mode that has both lifelong economic and cognitive consequences (Bialystok, 2009; Bialystok, Craik, & Freedman, 2007; Kovács & Mehler, 2009).

In this article we focus on a comparative investigation of monolingual and bilingual first language acquisition by asking how learning two languages at once during early childhood influences the rate and order of word learning. We investigate the rate by looking at both word and concept learning, two aspects of language learning that have often been difficult to discern between monolingual and bilingual populations because of small sample sizes. We investigate the order both by comparing two recent

* Data and code sufficient to reproduce the analyses and figures can be found here: www2.warwick.ac.uk/fac/sci/psych/people/thills/thills/bilsonetal_data.zip.

* Corresponding author at: Department of Psychology, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom. Tel.: +44 247 652 3183.

E-mail address: t.t.hills@warwick.ac.uk (T.T. Hills).

models of word acquisition and by asking what is the structure of bilingual children's growing semantic (specifically, associative) networks. This allows us to compose a quantitative picture of how two languages influence one another at the semantic level, either by inhibiting or facilitating the learning of words that are synonymous across languages—what are called *translational equivalents*.

Before describing our approach in more detail, we first provide the context for our investigation by taking a brief tour of what is known about cross-language influence and the comparative differences between monolingual and bilingual children during early language acquisition.

1.1. Cross-language interaction

The answer to the question of how languages influence one another during early development has evolved over the past century (Genesee, 2006). Theoretical positions on early language influence fall into roughly three camps. One of the earliest accounts—the *unitary account*—suggests that early language learners (prior to 24 months) learn languages as if they are learning a single undifferentiated language (Leopold, 1939; Redlinger & Park, 1980; Volterra & Taeschner, 1978). This was based on evidence that early bilinguals often show language acquisition patterns indicative of inhibition—words in one language appearing to slow the acquisition of words in the other. Volterra and Taeschner (1978), for example, demonstrated that among three bilingual children they studied the children tended to have a limited knowledge of translational equivalents. Researchers also observed that bilingual children often code-mix, using speech that contains both languages in a single phrase, or produce cross-linguistic blends (i.e., new words composed of two or more lexemes taken from different languages) such as the German/English word *pfeifing* to mean 'whistling' (Redlinger & Park, 1980). In sum, proponents of the unitary account hold that limited translational equivalents, code-mixing, and blends are the consequences of what initially starts out as one language, and only slowly differentiates into two languages.

In contrast, a second account—the *dual language account*—holds that bilingual first language learners learn their two languages independently from the earliest stages of word learning. This is not unreasonable since children who grow up in multilingual environments are able to differentiate between the languages they hear well before they begin language production (Bosch & Sebastián-Gallés, 2001; De Houwer, Bornstein, & De Coster, 2006; Junker & Stockman, 2002; Werker & Byers-Heinlein, 2008). By the dual language account, there is, thus, no influence between the two languages. Proponents of this account claim that code-mixing and blends are the result of simple pragmatic and sociolinguistic competence, not the result of grammatical or lexical confusion (Genesee, Nicoladis, & Paradis, 1995).

More recently, accounts based on language interaction—*emergentist accounts*—have become increasingly prominent. According to emergentist accounts, languages influence one another during development, for example, via processes involving competition for referents and

parasitic relationships where one word utilizes the conceptual packaging invoked by a word in the other language (Hernandez, Li, & MacWhinney, 2005; MacWhinney, 2004; Shook & Marian, 2013; Yip & Matthews, 2000). Further evidence also suggests that—among, for example, children of reading age engaged in acquiring L2—languages may interact via transfer of derivational morphology (Deacon, Wade-Woolley, & Kirby, 2007; Pasquarella, Chen, Lam, Luo, & Ramirez, 2011; Ramirez, Chen, & Pasquarella, 2013). Thus, emergentist accounts, as we define them here, offer numerous routes for language interaction, including both inhibition and facilitation. However, unlike the unitary account, emergentist accounts make no claim that the languages start out as a mosaic of two-languages-in-one.

These three accounts provide a theoretical starting point for investigating language interaction, but they have less to say about potential differences in the rate of word and concept learning or differential use of strategies that may influence their order of acquisition. Numerous comparative studies between bilingual and monolingual first language learners have pointed to both similarities and differences in these aspects of word learning (see Bialystock et al., 2009). For example, past research has found that bilinguals and monolinguals appear to learn words for concepts at approximately the same rate (De Houwer, Bornstein, & Putnick, 2013; Hoff et al., 2012; see also Cattani et al., 2014). Bilinguals and monolinguals also show similar capacities for mapping words to objects in the learning environment (Byers-Heinlein & Werker, 2013; Werker, Byers-Heinlein, & Fennell, 2009) and appear to learn their first words at approximately the same time (Vihman, Thierry, Lum, Keren-Portnoy, & Martin, 2007). Unfortunately, the difficulty of collecting bilingual data often limits inferences about the influence of early bilingualism to small samples. Small samples lead to underpowered studies that can mask the detection of true differences. This is well understood in research on early language acquisition and on bilingualism in particular (e.g., Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2012).

Critically, bilingual and monolingual first language learners do exhibit important differences. For example, bilinguals and monolinguals differ in the way they perceive and categorize phonemes (Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009). They also differ in their usage of known words to disambiguate the meaning of novel words—a process called mutual exclusivity (Byers-Heinlein & Werker, 2009; Byers-Heinlein & Werker, 2013; Houston-Price, Caloghris, & Raviglione, 2010). Bilingual infants tend to show either less reliance on mutual exclusivity or slower development of its use during early word learning (Clark, 2009; Markman & Wachtel, 1988). This may facilitate cross-language learning of synonyms. However, it may also inhibit learning within a language, where known words can facilitate the acquisition of novel words (Mather & Plunkett, 2010). That is, if monolinguals use mutual exclusivity to enhance their learning rate, and bilingual infants reduce their use of mutual exclusivity, this should reduce their learning rate within and possibly across languages.

The above findings suggest that though concept learning rates may be similar among bilinguals and monolinguals, the patterns of word acquisition may be unique. There are several reasons for this. The first rests on differences in the usage of mutual exclusivity, as described above. If bilingual children relax mutual exclusivity in their word learning, this may make learning of close associates more difficult, because mutual exclusivity may be unavailable as a means to differentiate things that appear together but have different meanings. Learning the word 'drink' after learning the word for 'water' should be more difficult to semantically differentiate if the child believes that they could mean the same thing. However, to the extent that mutual exclusivity is preserved, as the unitary language account may be taken to suggest, children should under-learn translational equivalents. Thus, differential usage of mutual exclusivity may influence both the rate and order of word learning.

A second reason why word order may be influenced for bilinguals rests on bilingual children's capacity to use more well known words in place of less well known words. Simply put, two sets of labels for objects in the world provides infants with plausible alternatives during production, and this may reduce the learning rate of non-dominant words. This is similar to the process of competition described by Hernandez et al. (2005). Producing one name is potentially sufficient to communicate the object of reference, a phenomenon potentially indicated by the prevalence of code-mixing (Deuchar & Quay, 1999; Genesee et al., 1995). If its use is systematic in favor of dominant words, this would inhibit the learning of translational equivalents.

In contrast to competition, learning the names of objects in one language may facilitate the learning of a second name in another language similar to the way that labels facilitate the learning of new conceptual categories (Fulkerson & Waxman, 2007; Waxman & Markow, 1995). Word learning in young children can reasonably be thought of as solving two problems: (1) the concept problem—learning to identify a conceptual category and (2) the naming problem—learning to map a specific word to that category. Learning the word for a concept in one language requires solving both problems. Learning an additional word for a concept only requires solving the naming problem. Indeed, a recent study (Hendrickson, Kachergis, Fausey, & Goldstone, 2014) found that one of the quickest ways to learn a name for an abstract category was to have already learned a different word for that category. Thus, the learning of a word can act as a label for a category that can then be reassigned to a new label—facilitating word learning in the second language. In contrast to mutual exclusivity, this process should lead to facilitated acquisition of translational equivalents across languages.

1.2. *The associative structure of the early lexicon*

We address the above issues by combining models of word learning with a large collection of bilingual and monolingual word learning data. With respect to understanding differences in word order, our research relies on

building associative networks to understand how the statistical structure of the lexicon is altered by the learning of a second language. Network analysis has been used successfully to understand semantic relationships in monolingual language acquisition (Beckage, Smith, & Hills, 2011) as well as within languages more generally (Arbesman, Strogatz, & Vitevitch, 2010; Ninio, 2006; Serrano, Flammini, & Menczer, 2009). A network can be constructed from language by allowing words to be nodes and connections between words (i.e., edges) to be based on specific relationship between words. Examples of these relationships include shared features (Hills, Maouene, Maouene, Sheya, & Smith, 2009a), co-occurrences in text (Hills, Maouene, Riordan, & Smith, 2010), phonological neighbors (Siew, 2013; Vitevitch, 2008) or free association norms (Hills, Maouene, Maouene, Sheya, & Smith, 2009b; Steyvers & Tenenbaum, 2005). These relationships allow one to formally describe statistical relationships between words as well as entire lexicons and to use these as predictors for when and how words will be added into a child's lexicon.

In the present work we use free association norms as a proxy for understanding semantic relationships between words. In previous work, we showed that the statistical structure of adult free associations could predict how monolingual children's early semantic networks grew between 15 and 25 months of age (Hills et al., 2009b). Moreover, the formal description of this model (called preferential acquisition) was a better predictor of word learning than a model based on the semantic structure of the words children already knew (preferential attachment). This work has since been corroborated with additional data sets and methods for establishing semantic similarity (Hills et al., 2010; Sailor, 2013). This indicates that, among other influences—such as phonological neighborhoods, word repetition, and pragmatics (Lust, 2006; Storkel, 2004; Tomasello, 2009)—one of the key influences on early word learning is the statistical structure of the language learning environment. Consistent with this claim is the observation that adults amplify the associative structure of their language when speaking to children, specifically producing more associates (as produced by adults in free association norms) around words that children learn earliest (Hills, 2012). In addition, several studies have now established that children during their earliest years of language learning are influenced by the semantic (i.e., associative) relations between words (Arias-Trejo & Plunkett, 2013; Delle Luche, Durrant, Floccia, & Plunkett, 2014; Styles & Plunkett, 2009; Willits, Wojcik, Seidenberg, & Saffran, 2013).

In sum, the network approach used here allows us to take a first step toward answering the question of how learning two languages at once influences the growth rate and statistical structure of the early lexicon. Moreover, the modeling approach also allows us to do something that has been difficult to do in the past: simulate the development of a bilingual lexicon as the product of two monolinguals. With this, we can further ask how the structure of the bilingual lexicon differs from that predicted by monolingual growth.

2. Methods

2.1. Vocabulary checklists

Data were collected from parents in association with child language studies conducted by the second author at the University of Houston. Caregivers were provided with a vocabulary checklist (two separate checklists in the case of bilinguals) and instructed to mark the words on each form that he/she had heard the child spontaneously produce (as opposed to merely imitating). Specifically, caregivers filled out the MacArthur–Bates Communicative Development Inventory (MCDI) Toddler Form for English words (Dale & Fenson, 1996; Fenson et al., 1994), or similar communicative development inventories for Spanish words (Jackson–Maldonado, Bates, & Thal, 1992), for Japanese words (Ogura & Watamaki, 1997), for Mandarin words (Tardif, Fletcher, Liang, & Zuo, 2008), or similar non-English versions constructed to resemble the MCDI but also take into account culture-specific words necessary for measuring vocabulary for specific child-language studies. The data used here were collected using a version of these forms that included approximately 430 items in each language. Animal sounds were excluded from the present analysis.¹

The MCDI is well suited for specific research questions regarding translational equivalents in bilingual preschool children because of its comprehensive nature, including a large number of items from a variety of lexical categories and a large number of overlapping concepts across languages. Moreover, the MCDI correlates positively and significantly with laboratory observations of vocabulary and has been shown to accurately reflect monolingual and bilingual language acquisition (Dale, 1991; De Houwer et al., 2006; Fenson et al., 1994; Marchman, 2002; Pearson, Fernández, & Oller, 1993; Poulin–Dubois et al., 2012).

For the purposes of the present study, we used all the children in the data described above for which we had complete entries and for which the child was either growing up in a monolingual English environment or growing up in a bilingual environment where one of the languages spoken was English. Because data were collected at several time points for some children, we always took the data collected at the earliest time point. Thus, this data consisted of 254 English-speaking monolinguals (129 females; mean age = 38.99 months, range = 6.15–78.39) and 181 bilinguals (89 females; mean age = 40.64 months, range = 6.28–92.14). Because we were interested in the general influence of learning a second language on a common language for which we could use the same methodology, we chose English as the common language because it was the most prevalent in the data and the most well studied with regard to associative structure. We included all additional languages in the data that met the above criteria regarding bilinguals. Languages spoken by bilinguals were as follows: Spanish ($n = 111$), Mandarin ($n = 25$), Vietnamese ($n = 19$),

Malayalam ($n = 9$), Japanese ($n = 8$), Arabic ($n = 4$), French ($n = 4$), and Russian ($n = 1$).²

There were 240 words whose concept translations appear in all the separate language checklists. We used these words to study the acquisition of translational equivalents. Thus, as in prior research on translational equivalents, we focused on the list of words shared by all languages in our study (De Houwer et al., 2006). Also, as noted in prior research (De Houwer et al., 2006), not all children exposed to bilingual input produce words in both languages. In those cases where we investigate the production of translational equivalents, we limit our investigation to those children producing words in both languages ($n = 105$).

2.2. Construction of networks

The words collected above provide the nodes in the child semantic networks we use to investigate lexical structure. The edges between nodes were constructed using the University of South Florida Free Association Norms (FAN; Nelson, McEvoy, & Schreiber, 2004). These norms were collected by providing a word (the cue) and asking an adult to provide a word—the first word that came to mind—in response (the target). From this process, one can establish cue–target pairs. For example, if the cue is ‘dog’, a participant might respond with ‘cat’ (the target). This constructs the associative pair dog–cat. The FAN consists of 5044 word cues and was used to construct an adjacency matrix F such that

$$F_{ij} = \begin{cases} 1 & \text{Word } j \text{ is the target of word } i \\ 0 & \text{Otherwise} \end{cases}$$

Throughout the rest of this work we define the associative indegree of word i , the number of cues that lead to the production of that word in the free association norms, as $k_i = \sum_{j=1}^n F_{ij}$, and the average degree as $\langle k \rangle = \sum_{i=1}^n k_i / n$. For those cases where we use the FAN, we used all 395 words from the MCDI that were also in the FAN.³

3. Results

3.1. Do bilinguals and monolinguals learn words or concepts at different rates?

We first consider the growth of bilingual and monolingual lexicons over time. To model the acquisition of words and concepts produced by monolingual or bilingual children we make use of a statistical model of language acquisition that treats all words equally—i.e., the binomial

¹ Less than 2% of the words were etymologically and orthographically identical between Spanish and English (e.g., ‘animal’). Removing these words from the analysis did not influence the results.

² Past research has often limited bilingual studies to the study of two specific languages and generalized to other bilinguals accordingly. The present approach trades off lexical specificity for lexical breadth by including all bilinguals for which we have available data. The results presented here are qualitatively similar if we limit the data to only Spanish–English bilinguals, or exclude all Spanish–English bilinguals. None of the conclusions that are drawn from the data are influenced by separating the data in this way. Thus, our results suggest general features of bilingual language acquisition.

³ Using associative strengths instead of presence or absence of an edge provided similar results. The computations are more straightforward using presence-or-absence and without loss of generality.

distribution. Let $x(t)$ be the number of words or concepts a child knows at month t , integer valued in the range $[0, N]$. The earliest record of word learning in our data for both monolingual and bilingual children was at month 6. Thus, we let $x(6) = 0$ and $\frac{dx}{dt}|_{t=6} = 0$. We also note that there is an upper bound, $x \leq N$, due to the checklist being a subset of the actual number of words/concepts produced by the child. Thus our model should show asymptotic behavior, $\lim_{t \rightarrow \infty} x(t) = N$.

We consider a two parameter binomial distribution showing such behavior, $x|t \sim B(N, q(t, s, \lambda))$, where the probability of knowing a word is modeled as a growth curve using the Weibull function with latency, λ , and shape parameter, s .

$$q(t, s, \lambda) = 1 - 2^{\left(-\frac{t-\lambda}{s}\right)^s}, \quad \lambda > 0, \quad s > 0 \quad (1)$$

When $s > 1$ learning accelerates over time. With $s < 1$ learning decelerates. When $s = 1$ the model reduces to the exponential. The month at which half the MCDI words are known is $\lambda + 6$. From this, we can construct the likelihood function $L(s, \lambda) = p(\mathbf{x}|\mathbf{t}, s, \lambda)$ given data $(\mathbf{x}, \mathbf{t}) = \{(x_i, t_i)\}_{i=1}^n$ for n children as follows:

$$L(s, \lambda) = \prod_{i=1}^n \binom{N}{x_i} q(t_i, \lambda, s)^{x_i} (1 - q(t_i, \lambda, s))^{N-x_i} \quad (2)$$

By the usual technique of maximizing the log-likelihood for both the monolingual and bilingual datasets, we find the maximum likelihood estimators (MLE) $\hat{\lambda}_M, \hat{s}_M$ and $\hat{\lambda}_B, \hat{s}_B$ for monolinguals and bilinguals, respectively. Additionally, we use a likelihood ratio test to determine whether or not the shape and scale parameters differ between monolinguals and bilinguals. This involved fixing the shape or scale parameter, while allowing the other parameter to vary, such that the likelihood ratio could be compared between statistical models with and without the specific shape or scale parameters for monolinguals and bilinguals (Lewandowsky & Farrell, 2010).

The resulting best fits and MLEs are given in Fig. 1.⁴ It is clear that, with respect to English words, monolingual English children learn words at a faster rate than bilingual children. The 95% confidence intervals for the number of unique words learned does not overlap for the two groups, and $\hat{\lambda}$ —the point at which children learn approximately half the words in the MCDI—is approximately 10 months earlier for monolinguals than bilinguals. Both groups show acceleration in their vocabulary growth over time, consistent with previous research (McMurray, 2007), but bilinguals show less acceleration than monolinguals ($\hat{s}_M = 1.64$ vs. $\hat{s}_B = 1.42$).

Concept learning does not show the same degree of difference but still reveals a lag for bilinguals. Again, the 95% confidence intervals do not overlap for the majority of the age range and $\hat{\lambda}$ shows an approximate 4.5-month advantage for monolinguals over bilinguals. However, the shape

parameter is approximately the same for concept learning between monolinguals and bilinguals. Likelihood ratio tests reveal that the differences in parameter values for monolinguals and bilinguals are warranted for all cases ($p < .001$) except for the shape parameter in concept learning, where both monolinguals and bilinguals show similar rates of acceleration. These results are consistent with the observation that the frequency of words in the learning environment is smaller per language for bilinguals, as well as directly correlated with vocabulary knowledge (David & Wei, 2008; Pearson, Fernández, & Oller, 1995; Pearson et al., 1993). However, our results further suggest that this difference may have an additional influence on the rate at which children acquire conceptual vocabulary. We discuss this further below.

3.2. Properties of earliest learned English words

For monolinguals, a word's age of acquisition is correlated with its associative relationships with other words (Hills, 2012; Hills et al., 2009b). If bilingual children show the same pattern, then their earliest learned English words should also show a bias for higher associative indegree. On the other hand, if the preference for words with higher associative indegree is diluted by a second language—created by competitive interference between the two languages—then we should expect the correlation between age of acquisition and associative relationships to be reduced for bilinguals. We define age of acquisition via logistic regression as the age at which 50% of children know the word. To compute indegree, we consider two network sizes using the final size child FAN network (395 words) and the full adult FAN network (5044 words). Fig. 2 shows the two measures of indegree in relation to age of acquisition (AoA) of English words for both monolingual and bilingual children. In both cases, the preference for words with more associative relationships is enhanced.

A regression analysis shows that, in both cases, bilinguals and monolinguals have different intercepts (reflecting the delayed English word learning in bilinguals) and have significant or marginally significant interactions with the log of the degree. The results of the regressions are as follows: for the child FAN network, the main effect of language, $\beta = 7.02$, $SE = 0.51$, $p < 0.001$; main effect of $\log(k)$, $\beta = -0.32$, $SE = 0.05$, $p < 0.001$; the interaction effect, $\beta = -0.14$, $SE = 0.07$, $p = .053$; and overall $R^2 = .32$. For the adult FAN network: main effect of language, $\beta = 7.17$, $SE = 0.52$, $p < .001$; the main effect of $\log(k)$, $\beta_M = -0.03$, $SE = 0.006$, $p < .001$, $\beta_B = -0.05$, $SE = 0.007$, $p < .001$; the interaction effect, $\beta = -0.02$, $SE = 0.01$, $p = 0.02$; and overall $R^2 = .30$. In sum, bilinguals and monolinguals show similar preferences for early learning of words with high associative indegree. However, the interaction effects indicate that the bilinguals, relative to monolinguals, show an enhanced preference for words with higher associative indegree.

3.3. Modeling growth of semantic networks

Previous work has investigated several learning rules associated with the growth of semantic networks based

⁴ For all statistical analyses in this article, additional analyses were run removing individuals who were at ceiling for the English MCDI, or who knew all translational equivalents. In all cases, the pattern of results was statistically and qualitatively similar to those reported in the text.

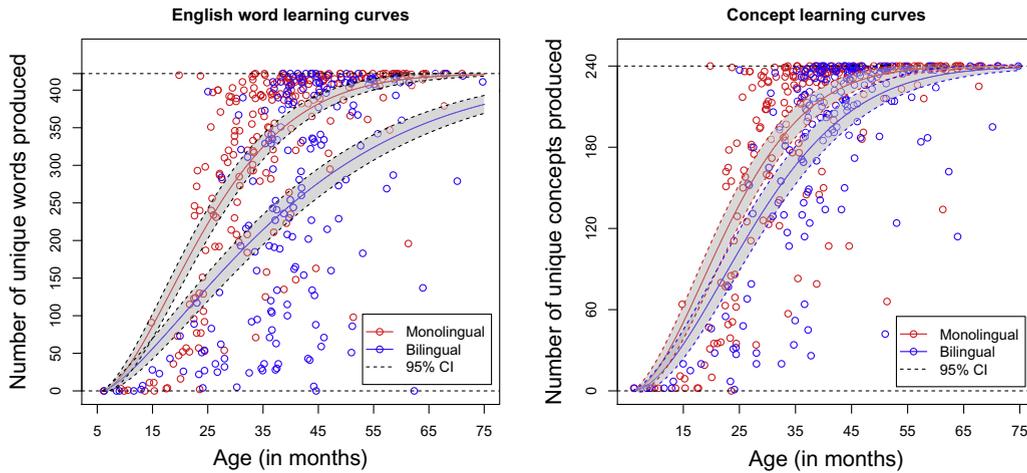


Fig. 1. Growth curves for word learning in monolingual and bilingual children. Left: ($\hat{\lambda}_M = 19.10$, 95% CI [18.94, 19.26] and $\hat{s}_M = 1.64$, 95% CI [1.62, 1.67] vs. $\hat{\lambda}_B = 29.63$, 95% CI [28.83, 29.64] and $\hat{s}_B = 1.42$, 95% CI [1.41, 1.49]). Right: $\hat{\lambda}_M = 17.82$, 95% CI [17.61, 18.03] and $\hat{s}_M = 1.78$, 95% CI [1.74, 1.81] vs. $\hat{\lambda}_B = 22.33$, 95% CI [22.04, 22.61], $\hat{s}_B = 1.80$, 95% CI [1.76, 1.84].

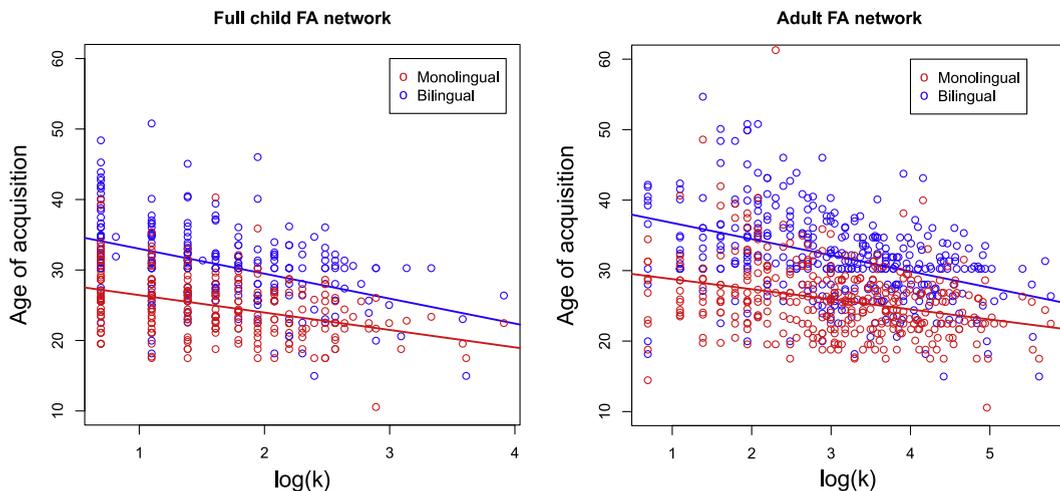


Fig. 2. The age of acquisition for monolinguals and bilinguals in relation to indegree computed for the child semantic network ($n = 395$ words) and the full adult free association norms ($n = 5044$ words).

on learning words in relation to associative indegree (Hills et al., 2009b). In this section, we examine two versions of these learning rules: preferential acquisition and preferential attachment. For preferential acquisition, words are learned in relation to their associative indegree, a relationship that is known to be more salient in the early learning environment as a result of changes in child-directed speech (Hills, 2012). An example of preferential acquisition would be learning the word ‘ball’ more quickly than ‘telephone’ because ‘ball’ is a word that is produced more often as a cue in free association norms. In preferential attachment—a network growth process that can also lead to higher indegree for words learned earlier—words are learned in relation to their capacity to attach to already well-connected words in the network. An example of preferential attachment would be learning the word

‘bump’ more quickly than the word ‘couch’ because ‘bump’ will connect to the already well-connected word ‘car,’ whereas ‘couch’ will connect only to the less well-connected word ‘chair.’

We consider two one-parameter models that show this behavior.⁵ Both models select a unique word (node i) from the subset W of the English MCDI word list that is present in the FAN data (395 words in total). It is then added to the existing network according to a probability distribution $P(i)$ dependent on the indegree K_i of that word (for preferential

⁵ Other models of preferential attachment and acquisition were considered with similar results. All correspond to the basic properties of attachment based on the existing structure of the network (preferential attachment) or acquisition based on the structure of the lexicon to be acquired (preferential acquisition).

acquisition) or a combination of the indegrees k_j in the child network given that word i was added and attaches to nodes j (for preferential attachment). Here we define K_i to be the number of unique cue words for which word i is a target in the entire adult FAN dataset of 5044 words. A realization of these models is constructed as follows: A word $i \in W$ is sampled without replacement and added to the child network according to the following discrete probability distributions

$$P(i) = \begin{cases} \frac{\sum_{i \in W} (K_i + 1)^\beta}{\sum_{i \in W} (K_i + 1)^\beta} & : \text{Acquisition} \\ \frac{\prod_{j_i} (k_j + 1)^\beta}{\sum_{i \in W} \prod_{j_i} (k_j + 1)^\beta} & : \text{Attachment} \end{cases} \quad (3)$$

with $\beta > 0$ in both cases. The next word is then sampled via Eq. (3) from the remaining word list. This process is repeated until no words remain. For every $N \in \{1, 2, \dots, N_{max}\}$, $N_{max} = 422$, a network was formed using the words selected by Eq. (3). This gives one realization of our network growth model. We then averaged over 500 realizations, calculating the mean growth curves of $\langle k \rangle(N)$. Our best-fit value $\hat{\beta}$ was then calculated for each model by minimizing the MSE (mean squared error) between our model and the dataset $\left\{ \langle k \rangle_j, N_j \right\}_{j=1}^n$, i.e.

$$\hat{\beta} = \arg \min_{\beta} \frac{1}{n} \sum_{j=1}^n \left(\langle k \rangle(N_j) - \langle k \rangle_j \right)^2 \quad (4)$$

The results are given in Table 1. Consistent with prior work on monolinguals, preferential acquisition performs considerably better for both monolingual and bilingual children. The $\hat{\beta}$ for bilinguals is also slightly higher than for monolinguals, consistent with the observation from Fig. 2 indicating the bilinguals show an enhanced preference for more associative words.

For further comparison, we present the results of the simulation in Fig. 3. We also include a simulation of random word learning, which samples from a uniform distribution of the word list in the same way as described above (i.e., $\beta = 0$). Plots of $\langle k \rangle$ along with $\pm 2\sigma$ prediction intervals are included for each N , where σ is the standard deviation from 500 realizations. The plots show a clear difference between random word learning and the data for monolinguals and bilinguals. Moreover, the plots show that preferential acquisition captures the main swathe of the data.

These results show that the general order in which monolinguals and bilinguals learn English words is very similar, favoring words with higher associative indegree. The larger MSE for preferential attachment along with the visual evidence that preferential attachment shows a different pattern of growth with respect to $\langle k \rangle$ is consistent with previous evidence that preferential attachment is of limited utility in its predictive accuracy regarding early word learning (Hills et al., 2009b) – though, in other contexts, such as language evolution, preferential attachment potentially has broad utility (Steyvers & Tenenbaum, 2005).

The above analyses provide some indication that second language learning may alter the learning trajectory for English words, even though bilingual and monolingual

Table 1

Mean squared error (MSE) and best fitting parameter values for the preferential acquisition and preferential attachment models.

	Acquisition	Attachment
MSE-Monolingual	0.017	0.057
MSE-Bilingual	0.049	0.159
$\hat{\beta}$ – Monolingual	0.017	0.31
$\hat{\beta}$ – Bilingual	0.019	0.31

acquisition patterns largely follow the same trajectory. However, because we have so far dealt with language in the aggregate, this tells us little about how the learning of a word in one language influences the learning of words in a second language. In the following sections we look at the structural differences in the lexicons of monolingual and bilingual first language learners, and then we address individual word learning.

3.4. Is English semantic network structure altered by bilingual first language acquisition?

Fig. 4 presents typical English language networks for monolingual and bilingual children, with edges defined by the free association norms. There are numerous network metrics available for analyzing the statistical structure of such networks. Here we use a subset of measures often associated with network analysis and indicative of general connectivity. These are the number of nodes (words), N , the clustering coefficient, C , the average path length, L , and the density of the network, ρ . In addition, we also compute a quantitative measure of small world structure, the small world index, S , which provides a graded measure of internal connectivity relative to what may be expected for random graphs of the same size (Humphries & Gurney, 2008). These network measures provide a blunt but common point of reference for evaluating how the semantic structure of learned words changes as a result of learning two languages at once.

The clustering coefficient, C , measures the extent to which neighbors of a node are connected. A common measure of clustering is computed based on the number of triplets of nodes that are fully connected (i.e., triangles). Because of the limited number of complete directed triangles in our networks, we use the undirected network to compute this measure. In the Monolingual English network in Fig. 4, *soft*, *nice*, and *pretty* in the upper middle of the graph representation, represent a complete triangle. The clustering coefficient is therefore the number of complete (undirected) triangles, φ , relative to the number of possible triangles, τ , such that

$$C = \frac{3\varphi}{\tau} \quad (5)$$

The average path length, L , represents the average over nodes of the shortest path length from a node to all other nodes in the network. The density, ρ , of the network represents the proportion of edges that are filled in the network.

Small world networks are defined as networks that have high clustering coefficients relative to a network of the same size and density but with randomly assigned

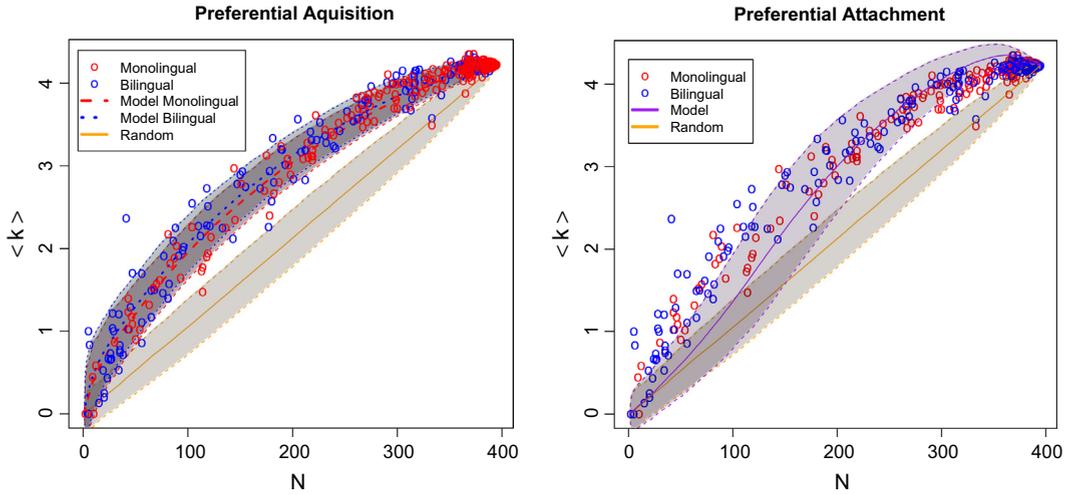
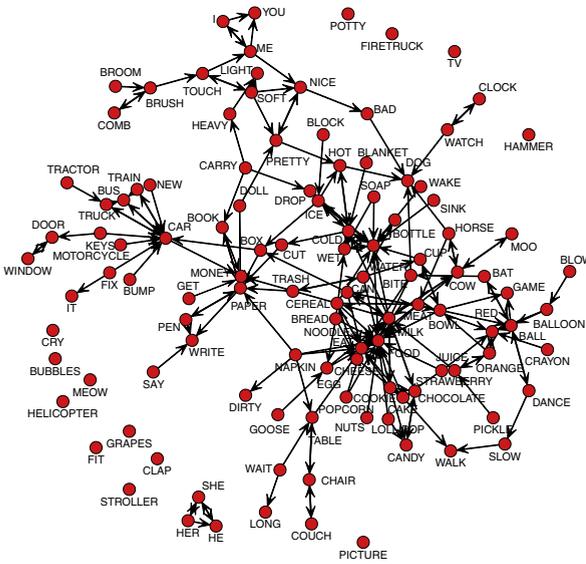


Fig. 3. Best-fit models of network growth using preferential acquisition, preferential attachment, and random learning.

A Monolingual English child FA network



A Bilingual (English/Spanish) child English FA network

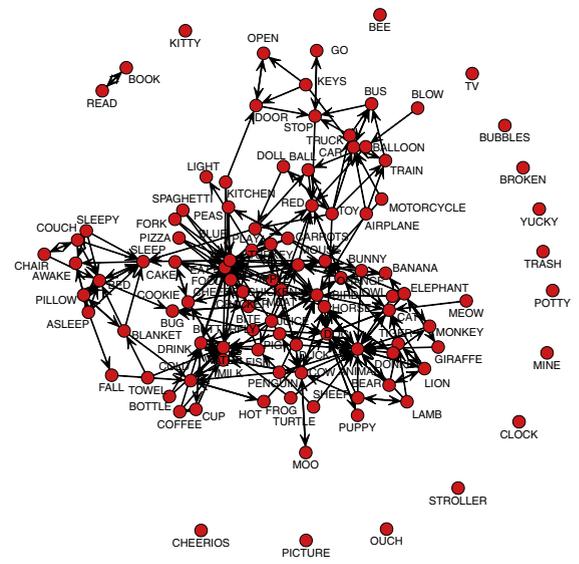


Fig. 4. Representative English networks for a monolingual and Spanish–English bilingual based on free associations (FA).

edges—commonly referred to as an Erdős-Renyi (ER) random graph (Watts & Strogatz, 1998). Not all natural networks are small worlds and small world properties have also been shown to differ for children who learn languages at different rates (Beckage et al., 2011). Here, we compute the small world index (Humphries & Gurney, 2008), based on the ratio of the observed clustering coefficient to the ER clustering coefficient, as

$$\gamma = \frac{C_g}{C_{ER}} \tag{6}$$

and the ratio of the observed average path length to the ER average path length,

$$\eta = \frac{L_g}{L_{ER}} \tag{7}$$

with the small world index,

$$S = \frac{\gamma}{\eta} \tag{8}$$

To compute S for each child, we randomly generated 500 ER graphs with matching numbers of nodes and density for each child. From these, we then took C_{ER} and L_{ER} as the mean clustering coefficient and average path length, respectively, from the size-matched random ER networks.

Table 2

Nonparametric regression table for properties of English networks for monolingual and bilingual children. k = average indegree, C = clustering coefficient, L = average path length, ρ = density, and S = small world index.

	Monolinguals	Bilinguals	F	RSS	p
k	3.58 ± 0.07	3.00 ± 0.10	17.81	14.05	<.001
C	.17 ± .003	.18 ± .008	0.83	1.89	.42
L	12.09 ± 0.20	11.03 ± 0.33	61.21	1336.50	<.001
ρ	.013 ± .0003	.018 ± .0017	2.03	0.076	.14
S	6.61 ± 0.07	6.25 ± 0.12	2.24	750.75	.002

Because the mean number of nodes is different for monolinguals ($M = 302.20$, $SE = 7.51$) and bilinguals ($M = 232.73$, $SE = 10.51$), $t(346.47) = 5.38$, $p < .001$, and the network statistics are constrained at the largest and smallest network sizes to be identical for monolinguals and bilinguals, we computed a nonparametric multiple regression controlling for the number of words in the network for each of the network measures.⁶

Table 2 presents the results of the analyses on these network measures. The results show that, after controlling for the number of words in the network, bilingual and monolingual children do not significantly differ on clustering coefficients or density. Moreover, via the small world index, both monolingual and bilingual semantic networks show a high-level of small-worldness, $S > 3$, on the order of that found for peer-to-peer networks and metabolic networks (Humphries & Gurney, 2008).

However, the structure of the English bilingual networks is influenced by learning a second language. The local polynomial regression fit for the significantly different variables in Table 2 are shown in Fig. 5, which provides a good indicator of the relative value of each of the parameters for different lexicon sizes. In particular, average indegree, k , and average path length, L , is for bilinguals larger than it is for monolinguals over a range of lexicon sizes. The small world index, on the other hand, is fairly indistinguishable during the earliest stages of word learning, but later becomes smaller for bilinguals than monolinguals. These results are consistent with the increased preference for higher indegree words among bilinguals, but also suggest that, like late talkers (Beckage et al., 2011), bilinguals may show an early preference within a language for words more distantly related to the words they already know. In the next sections, we investigate these findings directly in relation to the potential for facilitation or inhibition of translational equivalents.

3.5. Associative properties of translational equivalents

Here, we ask whether or not translational equivalents follow the predicted pattern of increased preference for words of high associative indegree. That is, if we imagine the word learning process in both monolinguals and bilinguals as a function of both a systematic learning process—

with a preference for words with high associative indegree—plus noise, then translational equivalents learned by bilingual children should reflect the overlap of two systematic processes with a shared preference for similar words. With respect to associative indegree, this means that words that a bilingual child learns first in both languages should have higher associative indegree than the English words (in the same child) that do not share a translational equivalent. If, on the other hand, words across languages inhibit one another, translational equivalents should not show a systematic property of having higher associative indegree than words with non-translational equivalents.

To investigate this, we compared the associative indegree of words that either were or were not translational equivalents in bilingual children's lexicons. As shown in Fig. 6, the bilingual children's translational equivalents have a higher associative indegree than their non-translational equivalents (results of a paired t -test: $t(215.7) = 17.42$, $p < 0.001$). These results support the hypothesis that bilinguals show shared preferences for word acquisition in both languages, because they are learning words with high associative indegree in both languages at the same time. This may also reflect a facilitative process, one that amplifies the acquisition of high indegree words in both languages.

3.6. Is learning translational equivalents more than a shared preference for the same words?

The above analyses suggest a shared preference for similar words in both languages, but also reveal a tendency to alter the order of language acquisition when learning a second language at the same time. In particular, the results on associative indegree and age of acquisition suggest that bilinguals may enhance their preference for words with higher associative indegree. This could be the result of a facilitation process, by which children tend to favor words with high associative indegree in both languages, but following the learning of one of these words this further facilitates the learning of the same word in the other language.

To test this, we compared the observed fraction of translational equivalents in each bilingual child's lexicon to the expected fraction of translational equivalents that child would have if they were independently acquiring two monolingual English lexicons at the same time. We achieve this by modeling lexical acquisition in each child using preferential acquisition. Specifically, we use the value of $\hat{\beta}$ for monolingual word acquisition fit only to that portion of the English lexicon for which we had translational equivalents in all 10 languages ($n = 243$ words). This allows us to model the learning of potential translational equivalents in bilinguals.⁷ We also use the monolingual $\hat{\beta}$ because we want to compare whether or not

⁶ Nonparametric regressions were computed using the loess function in R with leave-one-out cross-validation used to fit the span measure. Statistical significance is reported via the F -test for the change in the residual sum of squares between models with and without the network measure of interest, but always including the network size.

⁷ If instead we simulated the growth of the full English lexicon ($n = 429$) we would necessarily overestimate the number of expected translational equivalents. This is because two lexicons selected from two different languages will not have the same number of potential cross-language synonyms as two lexicons both selected from English.

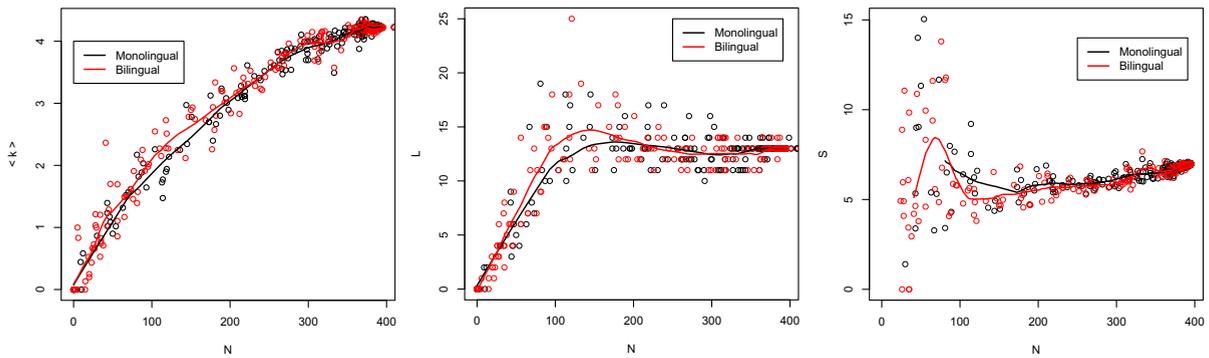


Fig. 5. Results of the nonparametric regression analyses showing the local polynomial regressions best fit to the data for average indegree, k , average path length, L , and small world index, S .

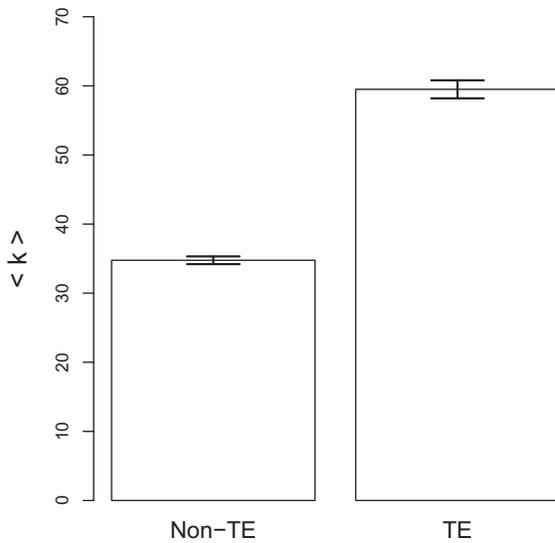


Fig. 6. Properties of translational equivalents in bilinguals. The mean associative indegree of words taken from the adult FAN for translational equivalents (TE) and non-translational equivalents (non-TE).

bilinguals show more or fewer translational equivalents than we would expect if we compared the lexical overlap of two randomly chosen monolinguals.

We then simulated the acquisition of 10^3 pairs of lexicons for each bilingual child, with the two simulated lexicons matched to the sizes of the bilingual child's lexicons. From this we computed the mean fraction of translational equivalents and the standard deviation for the lexicon sizes of the simulated bilingual dataset. A Z-score was then calculated for each bilingual child, based on the standard deviation and mean of the simulated data for that child alone. The results are presented in Fig. 7.

The data show that 24 bilingual children (out of 105) produce significantly more translational equivalents than expected if they were learning the two languages independently, and 9 bilingual children show significantly fewer translational equivalents than expected. The distribution of Z-scores for the bilingual children is significantly above zero ($M = 0.75$, 95% CI [.39, 1.11], $p < .001$), indicating that the majority of bilinguals overproduce translational

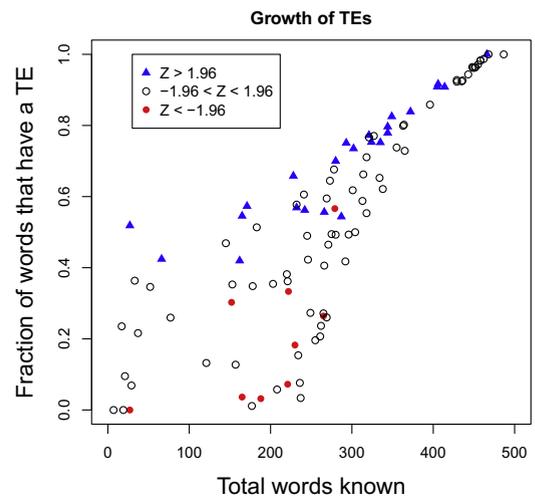


Fig. 7. Comparison of observed versus expected numbers of translational equivalents (TE) in bilingual children. The symbols provide an indication of whether or not the proportion of translational equivalents is significantly more or less than predicted by 10^3 pairs of simulated monolinguals with matched vocabulary sizes.

equivalents compared to their simulated means. This result is also found if we limit the data to only Spanish–English bilinguals ($M = 0.54$, 95% CI [0.06, 1.00], $p = .025$) or include only non-Spanish-speaking bilinguals ($M = 0.98$, 95% CI [0.41, 1.54], $p = .001$).⁸ These results suggest that word learning in bilingual children is facilitated by the learning of a second language, such that the children show a preference for learning words that they already know in the other language.

⁸ If we use the best-fit $\hat{\beta}$ for bilinguals, we are asking whether or not bilinguals show more or fewer translational equivalents than we would expect by comparing them with two English lexicons taken from different bilinguals. It turns out that the pattern of results from this analysis are quite similar to that for using the $\hat{\beta}$ taken from monolinguals, with 24 children showing significantly more translational equivalents than we would expect and only 10 showing fewer. A t -test on the Z-scores also shows a significant bias toward more translational equivalents ($M = .65$, 95% CI [0.29, 1.00]). This is a fairly strong test of language interaction, as it indicates that bilinguals' languages are more similar within bilinguals than across bilinguals.

4. Discussion

When understanding how children learn one language, we tend to focus on how the language learning environment shapes early learning. For bilingual first language learners, part of the language environment is represented by knowledge of a second language. This knowledge could be inhibitory, as one might speculate based on the principle of mutual exclusivity (Clark, 2009; Markman & Wachtel, 1988), or it might be facilitatory, with words in one language helping carve up the conceptual world for a second language. The picture we provide here favors the latter interpretation but also suggests that the relationship is still more complex.

First, though previous work with smaller collections of bilinguals has suggested that bilinguals learn concepts at a similar rate to monolinguals, our results suggest that—based on productive vocabularies—monolinguals outpace bilinguals during their earliest months of concept learning. This has many potential causes. If hearing a specific word in combination with a pattern of perceptual stimuli makes it easier to later identify that pattern as an entity that can be named, then bilinguals may be expected to show some degree of delay in concept learning. The frequency with which they hear a specific word in a specific context is likely to be smaller than a child who hears only one language. Alternatively, children may have learned the concept but may not have mapped it onto a specific word in either language. They may recognize the pattern as a thing that can be named, but due to a reduced frequency of co-occurrence with a specific word, they have more difficulty recalling a name for the concept. Studies looking at the difference between comprehension and production—well known to be associated with distinct learning processes (Bates, Dale, & Thal, 1995)—may be particularly relevant here. That is, the productive delay of concepts may not rely on concept knowledge alone, but may also be influenced by word-object mapping.

Second, our results are consistent with much prior work indicating that bilinguals and monolinguals follow the same path in language learning. Both groups show a clear affinity for learning hubs in the free association norms and follow a pattern of acquiring high associative indegree words early on. This is potentially a result of the architecture of child-directed speech, where the probability of producing associates around early learned words is nearly double that in adult-directed speech—thereby amplifying the semantic relations most around the words that children learn earliest (Hills, 2012). Associative structure may also be correlated with other properties of child-directed speech likely to influence early learning—such as repetitions, contextual diversity, and frequency (e.g., Hills, 2012; Hills et al., 2010)—and many other factors besides language structure are known to influence language acquisition (Lust, 2006; Tomasello, 2009). Thus, our analysis is not meant to be a claim that associative structure is the most important factor—or even that children are aware of this factor (Hills et al., 2010)—but simply that it is a factor indicative of early word learning, and one readily suited to the task of understanding early lexical

structure. Given the flesh-and-blood nature of the real-world process of language learning, it is perhaps impossible to ever show that a factor is completely irrelevant to learners who have access to that factor. Nonetheless, our approach does provide grounds for suggesting that monolinguals and bilinguals tend to grow their lexicons with a shared preference for words that have high associative indegree. Indeed, the enhanced associative indegree of translational equivalents learned by bilinguals is further evidence that bilinguals tend to value the same semantic properties of words in both languages.

Finally, our results contribute to a growing understanding of the role of language interaction in bilingual language acquisition (Deacon et al., 2007; Hernandez et al., 2005; MacWhinney, 2004; Pasquarella et al., 2011; Ramírez et al., 2013; Shook & Marian, 2013; Yip & Matthews, 2000). Specifically, learning two languages in one person is not simply a delayed version of learning two languages in two separate people. Knowledge of one language influences how a second language is learned. Specifically, our results suggest that one aspect of this interaction involves facilitating the acquisition of translational equivalents. Evidence for this is partly provided by differences in the statistical structure of the English semantic networks. The semantic networks' most prominent differences were found in average path length, average indegree, and their small world index. These are core features of networks and provide a common basis for structural comparison. Importantly, the results did not show the same pattern of low clustering coefficient as was previously found for late talkers (Beckage et al., 2011), but the reduced small world index and increased average path length amongst bilinguals is potentially symptomatic of a similar underlying process. That is, in addition to showing a preference for learning translational equivalents, bilinguals also show an early tendency to learn words that are more associatively distant in their English lexicon. This may have a simple explanation: bilinguals often speak different languages with different people in different contexts (Grosjean, 2010), and this may lead their lexicons in a given language to be more sparse relative to that of a monolingual. In other words, their early lexicons may lack certain words that act as intermediaries in early monolingual lexicons—with the consequence that their networks have larger average path lengths.

What is less consistent with the picture of context-specific language acquisition is that bilinguals show increased average indegree and simultaneously appear to show a preference for learning words with higher-associative indegree than monolinguals. This picture was made clearer by our computational model of bilingual language acquisition as two independent monolinguals. This showed that the majority of bilinguals tend to overproduce translational equivalents when compared with their vocabulary-size matched simulated monolinguals. Collectively, these analyses speak strongly in favor of facilitation. Perhaps as labels facilitate learning (Waxman & Markow, 1995) and are easily replaced with new labels (Hendrickson et al., 2014), naming something in one language facilitates the acquisition of a new word for that concept in another language.

The present work is not meant to be definitive but rather aims to provide some computational leverage on long-standing questions in bilingual first language acquisition. This follows an important tradition in computational approaches to understanding bilingual language acquisition (Li, 2013; Li & Farkas, 2002; MacWhinney, 2004). Nonetheless, our work is subject to methodological constraints. Studies of child language acquisition are often limited by parental report measures and bilingual studies are likely to be particularly influenced (De Houwer et al., 2013). This is especially true in cases where caregivers may lack proficiency in one language or another. Language checklists are also unlikely to include all the words children know. Moreover, as we note above, comprehension and production are not the same thing. As such, our results should be understood with these caveats in mind.

In conclusion, our results indicate that bilingual first language learners do not learn two languages as if they were one, nor do they learn two languages independently. The languages influence one another, and alter both rate and order of acquisition. Their production of translational equivalents clearly indicates what others have found before, that children do not use a 'one form, one meaning' assumption (De Houwer et al., 2013) or show constraints one might predict based on mutual exclusivity. On the contrary, our work provides evidence that the languages interact in an emergent and symbiotic relationship, with words learned in one language increasing the likelihood of learning synonyms in another language.

Acknowledgement

This work was funded by the National Institute of Health grant (0R01HD058620) and the Foundation for Child Development (Young Scholars Program) awarded to the second author (HY) and a Mid-Career Fellowship from the British Academy to the last author (TH).

References

- Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010). Comparative analysis of networks of phonologically similar words in English and Spanish. *Entropy*, *12*, 327–337.
- Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, *128*, 214–227.
- Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 96–151). Oxford: Blackwell.
- Beckage, N., Smith, L. B., & Hills, T. T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE*, *6*, e19348.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*, 3–11.
- Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*, 459–464.
- Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*, 89–129.
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*, 29–49.
- Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, *12*, 815–823.
- Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, *128*, 407–416.
- Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreco, F., Dennis, I., et al. (2014). How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests? *International Journal of Language & Communication Disorders*, *49*, 649–671.
- Clark, E. V. (2009). On the pragmatics of contrast. *Journal of Child Language*, *17*, 417–431.
- Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research*, *34*, 565–571.
- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*, 125–127.
- David, A., & Wei, L. (2008). Individual differences in the lexical development of French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, *11*, 598–618.
- De Houwer, A., Bornstein, M. H., & De Coster, S. (2006). Early understanding of two words for the same thing: A CDI study of lexical comprehension in infant bilinguals. *International Journal of Bilingualism*, *10*, 331–347.
- De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2013). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*, 1–23.
- Deacon, S. H., Wade-Woolley, L., & Kirby, J. (2007). Crossover: The role of morphological awareness in French immersion children's reading. *Developmental Psychology*, *43*, 732–746.
- Delle Luche, C., Durrant, S., Flocia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science*, *17*, 948–955.
- Deuchar, M., & Quay, S. (1999). Language choice in the earliest utterances: A case study with methodological implications. *Journal of Child Language*, *26*, 461–475.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*, 5 [Serial No. 242].
- Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition*, *105*, 218–228.
- Genesee, F. (2006). Bilingual first language acquisition in perspective. In P. McCardle & E. Hoff (Eds.), *Childhood bilingualism: Research on infancy through school age* (pp. 45–67). Clevedon: Multilingual Matters.
- Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*, 611–631.
- Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard University Press.
- Hendrickson, A. T., Kachergis, G., Fausey, C. M., & Goldstone, R. L. (2014). Re-learning labeled categories reveals structured representations. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*, 220–225.
- Hills, T. T. (2012). The company that words keep: Comparing the statistical structure of child-versus adult-directed language. *Journal of Child Language*, *40*, 1–19.
- Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009a). Categorical structure among shared features in networks of early-learned nouns. *Cognition*, *112*, 381–396.
- Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009b). Longitudinal analysis of early semantic networks. *Psychological Science*, *20*, 729–739.
- 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*, 259–273.
- Hoff, E., Core, C., Place, S., Ruminche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*, 1–27.
- Houston-Price, C., Caloghris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*, 125–150.
- Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS ONE*, *3*, e0002051.
- Jackson-Maldonado, D., Bates, E., & Thal, D. J. (1992). *Inventario de Desarrollo de Habilidades Comunicativas [MacArthur Communicative Development Inventories (CDI)]*. San Diego: Singular.

- Junker, D. A., & Stockman, I. J. (2002). Expressive vocabulary of German–English bilingual toddlers. *American Journal of Speech-Language Pathology*, 11, 381–394.
- Kovács, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, 106, 6556–6560.
- Lanza, E. (2000). Language contact: A dilemma for the bilingual child or for the linguist. In S. Döpke (Ed.), *Cross-linguistic structures in simultaneous bilingualism* (pp. 227–245). Amsterdam: John Benjamins Publishing.
- Leopold, W. (1939). *Speech development of a bilingual child: A linguist's record: Vol. 1. Vocabulary growth in the first two years*. Evanston, IL: Northwestern University Press.
- Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage Publications.
- Li, P. (2013). Computational modeling of bilingualism: How can models tell us more about the bilingual mind? *Bilingualism: Language and Cognition*, 16, 241–245.
- Li, P., & Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. *Advances in Psychology*, 134, 59–85.
- Lust, B. (2006). *Child language*. Cambridge: Cambridge University Press.
- MacWhinney, B. (2004). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism* (pp. 49–67). Oxford: Oxford University Press.
- Marchman, V. A. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research*, 45, 983–997.
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157.
- Mather, E., & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, 105, 232–242.
- McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, 317 [631–631].
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, 36, 402–407.
- Ninio, A. (2006). *Language and the learning curve: A new theory of syntactic development*. Oxford: Oxford University Press.
- Ogura, T., & Watamaki, T. (1997). *Japanese communicative developmental inventories: User's guide and technical manual*. San Diego, CA: Singular.
- Pasquarella, A., Chen, X., Lam, K., Luo, Y. C., & Ramirez, G. (2011). Cross-language transfer of morphological awareness in Chinese–English bilinguals. *Journal of Research in Reading*, 34, 23–42.
- Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, 43, 93–120.
- Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language*, 22 [345–345].
- Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2012). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, 17, 57–70.
- Ramírez, G., Chen, X., & Pasquarella, A. (2013). Cross-linguistic transfer of morphological awareness in Spanish-speaking English language learners: The facilitating effect of cognate knowledge. *Topics in Language Disorders*, 33, 73–92.
- Ramon-Casas, M., Swingle, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, 59, 96–121.
- Redlinger, W. E., & Park, T. (1980). Language mixing in young bilinguals. *Journal of Child Language*, 7, 337–352.
- Sailor, K. (2013). Is vocabulary growth influenced by the relations among words in a language learner's vocabulary? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1657–1662.
- Serrano, M. Á., Flammini, A., & Menczer, F. (2009). Modeling statistical properties of written text. *PLoS ONE*, 4, e5372.
- Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition*, 16, 304–324.
- Siew, C. S. (2013). Community structure in the phonological network. *Frontiers in Psychology*, 4, 553.
- Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29, 41–78.
- Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, 25, 201–221.
- Styles, S. J., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, 1, 1–24.
- Tardif, T., Fletcher, P., Liang, W. L., & Zuo, Q. H. (2008). *The Chinese communicative development inventory (Putonghua and Cantonese versions): Manual, forms, and norms*. Peking University: Medical Press.
- Tomasello, M. (2009). *Constructing a language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.
- UNESCO (2003). *Education in a multilingual world, education position paper. Place de Fontenoy – 75007 Paris (France)*.
- Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh, and English–Welsh bilingual infants. *Applied Psycholinguistics*, 28, 475–493.
- Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language and Hearing Research*, 51, 408.
- Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, 5, 311–326.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442.
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology*, 29, 257–302.
- Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences*, 12, 144–151.
- Werker, J. F., Byers-Heinlein, K., & Fennell, C. T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 3649–3663.
- Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, 18, 1053–1075.
- Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism Language and Cognition*, 3, 193–208.