Network Analysis of a Large Sample of Typical and Late Talkers

Eva Jimenez (E.Jimenez-Mesa@warwick.ac.uk)
Department of Psychology, University of Warwick, University Road, Coventry,
CV4 7AL, United Kingdom
Thomas Hills (T.T.Hills@warwick.ac.uk)
Department of Psychology, University of Warwick, University Road, Coventry,
CV4 7AL, United Kingdom

Abstract

The focus of this paper is to examine differences in semantic network structure of late talkers and typical talkers to elucidate potential learning strategies used by late talking children. To address this question, we conducted network analysis on the vocabularies of 2,912 children, with 566 of those being late talkers. Contrary to previously reported findings, the results show that late talkers have well-connected vocabularies as measured by median degree, clustering coefficient, and mean distance, with more well-connected networks in some cases than their typical talking peers. Further analysis of word order suggests that late talkers may be selecting based on frequency and connectivity of the words in the learning environment, more so than typical talkers. The language processing difficulties in late talkers appear not to be associated with their semantic network properties. In sum, late talkers may initially benefit from using word frequency and word connectivity strategies to build well-connected vocabularies.

Keywords: semantic networks; network analysis; corpus analysis; language acquisition; late talkers; word frequency

Introduction

Children start learning words within the first and second year of life. Some children are slower than others, and some of the slowest children go on to show lifelong learning difficulties. These children, called late talkers, have been the subject of extensive research, both to understand how they learn but also to understand how their learning might be better facilitated to prevent lifelong problems. One of the outstanding questions in late talker research is to what extent late talkers are simply 'slower' versions of typical talkers? The alternative is that they show different learning strategies and therefore not only learn words more slowly but learn different words. In the present work we address this question using network analysis on the largest currently available sample of children's vocabularies. The aim is that by identifying similarities and differences in vocabulary acquisition, we can better identify the strategies that late talkers might be using, if indeed they are using strategies different from typical talkers. Before we explain our methods, we first briefly review the literature on network analysis and late talkers.

Network Analysis in Language Acquisition

Semantic network analysis has allowed researchers to explore language processing in adults (Wachs-Lopes & Rodrigues, 2016) and language acquisition in children (Hills, Maouene, Riordan & Smith, 2010; Hills, Maouene, Maouene, Sheya & Smith, 2009). In network analysis, also known as graph theory, words are modelled as vertices and relationships between words are modelled as edges. Semantic relatedness amongst words is the focus of the present study, however, other relationships have been used in the past, including features, phonology, and free associations (e.g., Li , Farkas & MacWhinney, 2004, Hills, Maouene, Maouene, Sheya & Smith, 2009).

According to Watts & Strogatz (1988) small world networks are "highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs". Small-world properties have been reported not only in adult vocabulary but also in toddlers as young as 15 months (Beckage, Smith, & Hills, 2011). Local structure, where words are connected in clusters, may represent semantic categories (Hills, Maouene, Maouene, Sheya & Smith, 2009b). Words between clusters may facilitate transitions between clusters, and therefore are believed to be critical in language processing (Cancho & Solé, 2001; Banavar, Maritan & Rinaldo, 1999).

Networks statistics can be computed to evaluate the connectivity of a lexical network. Three of them are considered in this work. The *degree* of a word is the number of ties that word has with other words. Calculating the mean or median of this measure provides the overall level of cohesion in a network. The *clustering coefficient* explores the degree of clustering of nodes within a network. Finally, the *mean distance* of a network shows the average of the shortest paths between all pairs of words which gives an idea of its global access. Many studies have included these statistics in their analysis to examine research questions related to language acquisition, e.g. studies on lexical growth (Hills, Maouene, Riordan & Smith, 2010), categorical memberships in the lexicon (Hills, Maouene, Maouene, Sheya & Smith, 2009b), and the influence of bilingual first-language learning on early English acquisition (Bilson, Yoshida, Tran, Woods & Hills, 2015).

Language Acquisition in Typical and Late Talkers

Extensive research has already taken place around those children with small vocabularies compared to their normed peers (for a review, see Desmarais, Sylvestre, Meyer, Bairati & Rouleau, 2008). Whereas many late talkers catch up with their peers in word production (so called 'late bloomers'; Thal, Tobias & Morrison,1991), some others will have language difficulties that drag on, to be later diagnosed with Specific Language Impairment (SLI) (Leonard, 2000). For the latter group, language problems will continue with comprehension, production and/or pragmatics (Leonard, 2000). Despite late bloomers' vocabulary improvement, they may still be more likely to experience difficulties in language-related tasks, such as reading (Rescorla, 2009). Defining and understanding the characteristics and strategies of late talker's vocabulary could help to develop effective early interventions.

The starting point for our research is Beckage, Smith and Hills (2011). In their study, network analysis was used to characterize the vocabularies of 66 typical and late talking children. Semantic relatedness of words, computed from word co-occurrence derived from the CHILDES database (MacWhinney, 2000), was used to connect the words in the child's vocabulary. Results showed that both typical and late talkers exhibit small-world structure, although late talkers present this to a lesser degree. The study suggests the existence of a relationship between the child's rate of lexical development and the connectivity of her individual network. This finding led the authors to hypothesize the possibility of an 'oddball' strategy used by late talkers: a preference to learn words that have lower semantic relatedness with words they already know. Thus, late talkers may use different learning strategies or have differences in their ability to discriminate word referents.

A later study by Nematzadeh, Fazly & Stevenson (2014) challenged these results. By means of computational modelling, the authors simulated typical and late talking word learners to explore differences in their semantic networks. Surprisingly, neither type of simulated word learner showed a small-world structure. Referring to Beckage, Smith and Hills' work (2011), the authors questioned the use of the same edges to link words of the networks of both typical and late talkers as it assumes that both groups learn the same knowledge about words. Moreover, the authors called into question the 'oddball strategy' alleging that late talking children do not possess enough information about the words to discern similarities and dissimilarities between words. However, their methodological differences may also explain the differences in the results. Whereas in Beckage, Smith and Hills' study (2011) the co-occurrence of words in child-directed speech generated semantic relatedness between words, Nematzadeh, Fazly & Stevenson (2014) used associative semantic information provided by a custom lexicon to link words with similar meanings.

Word frequency has also been investigated as one of the main influences in word learning. In Stokes (2010), two-year old typical talking children tended to acquire more high frequency words than late talking children, who in turn learned more words with a higher phonological neighborhood density. However, the question remains whether there is a difference between the two groups with respect to preferences for acquiring some words earlier than others, which may have given a different perspective of the learning strategies used by each group.

The present work further investigates the difference between early and late talkers using vocabularies taken from the open repository website *Wordbank* (Frank, Braginsky, Yurovsky & Marchman, in press), providing a sample of 2,912 children, of which 566 are considered late talkers. The methodology followed and the words selected are identical to those in Beckage, Smith and Hills (2011), the only differences being the number of children and the diversity of their backgrounds as they come from nine different American studies. Our principle question is, how does the semantic vocabulary structure differ between typical and late talkers? We ask two additional questions: (1) what is the relationship between word frequency and order of word learning for the two groups, and (2) what is the relationship between word frequency in the language learning environment and the connectivity between the words. These allow us to investigate additional pathways for language learning in late talkers.

Methods

Vocabulary

Publicly available vocabulary data for 5,450 children aged 16 to 30 months was downloaded on October 2016 from *Wordbank* (Frank et al., in press). Data is contributed by various researchers using the MacArthur-Bates Communicative Development Inventory (MCDI) (Fenson, Dale & Reznick, 1993). The data set used in this work was downloaded by selecting 'Words & Sentences' under forms and 'English' under language. To facilitate comparisons between late and typical talkers, we limited vocabulary sizes to between 20 and 220 words, a range where typical and late talkers overlapped that also allowed for meaningful network statistics. Few late talkers had a productive vocabulary size greater than 220 words. After limiting the vocabulary size, the final total number of children remaining was 2,912. Of the 2,912 children (aged 15 to 30 months), 566 have a vocabulary size atypical for their age and 2,346 presented a normal vocabulary size for their age. Late talkers were at or under the 20th percentile of their same-age peers. To calculate this, each child was assigned to a decile grouping according to their age and vocabulary size reported in the MCDI instrument. The decile grouping was looked up from a table of estimated percentiles on the *Wordbank* website created using a quartile regression with monotonic polynomial spline as the base function. Although the total number of words that can be recorded in the MCDI questionnaire is 680, in Beckage, Smith and Hills' study (2011) only 291 words were used which appeared on both the toddler and infant forms, allowing comparison across ages. The same words were selected in this study. The 291 words consist of 207 nouns, 50 verbs, 14 adjectives, 12 pronouns, 6 adverbs, 1 quantifier and 1 demonstrative. All categories in the MCDI were included except for 'Sound Effects and Animal Sounds', 'Helping Verbs', and 'Connecting Words'.

Semantic Relationships Between the Words

To link the words in each child's productive vocabulary, semantic relatedness between words was computed using co-occurrence statistics derived from an American English corpus of child-directed speech, CHILDES (MacWhinney, 2000). A surface proximity approach (see Evert, 2008) was used to determine the frequency in which each distinct word (node) in the corpus co-occurred with other words (collocates). An empty co-occurrences matrix was created and then populated by moving a window of span size 5 words forward through the corpus. As co-occurrences were encountered the count for that pair was incremented. A subset of this large matrix was created where the rows and columns intersected with the 291 words selected from the MCDI forms. Finally, the count values were converted to a simple binary representation.

Table 1: Distribution of the children's vocabulary and mean age (in parenthesis)

	Vocabulary size									
	20-40	40-60	60-80	80-100	100-120	120-140	140-160	160-180	180-200	200-220
LT	133 (23)	70 (24.5)	74 (25)	60 (26.1)	55 (27.2)	46 (27.7)	36 (28.3)	46 (28.5)	25 (29.2)	21 (29.9)
TT	427 (17.1)	304 (17.9)	213 (18.5)	197 (19.6)	191 (20.4)	159 (20.8)	187 (21.4)	190 (22.7)	205 (23.4)	262 (24.5)
Total	560	374	287	257	246	205	223	236	230	283

Random Acquisition Networks

In order to compare each individual child's network with similar size networks, 300 random acquisition networks were generated for each child. These networks have the same number of words *n* as the child's network, but the words are selected randomly from the set of 291 words. Then, each random network was linked using the values from the CHILDES matrix explained above. The same statistical properties computed for each child's vocabulary network were also computed on the 300 random acquisition networks and then averaged. These random network statistics provide the structure inherent in the language context without including the particular word learning pattern of the child, thus providing a point of comparison for each child's network, allowing us to compare children with different size vocabularies against a 'random' learner.

Word Frequency and Connectivity

Preferences for learning certain words over others was assessed with respect to vocabulary size. First, the sample was divided into two groups: late talkers and typical talkers. Then each group was ordered by increasing vocabulary size, creating subgroups. Within each subgroup and for each word a count was made of children that produced the word. Words were ranked based on their count. Differences in ranking between late and typical talkers was calculated by subtracting the respective ranking value for each word, allowing us to identify differences between the groups in their preference for learning certain words.

Only words with a minimum ranking difference of 20 or more are presented, but the results are not sensitive to this number. The frequency of each word was taken from CHILDES and compared between the two groups (see MacWhinney, 2000; Li & Shirai, 2000). Within the 291 x 291 matrix of co-occurrences, some words are better connected (have higher degree) than others. This was calculated directly from the matrix by counting the total number of occurrences of a word with other words.

Results

Network Analysis

Analysis was carried out using R and the *igraph* package, version 1.0.1 (Csárdi & Nepusz, 2006). Connectivity was assessed by computing three statistics of each directed network: median in-degree, clustering coefficient, and mean distance. Late talkers are unequally distributed across the sample: they have higher representation at lower vocabulary sizes and low or no representation at the higher vocabulary sizes. Therefore, the sample was divided into bins of children with similar vocabulary sizes in ranges of 20 words. The size of the bin does not influence our results, but does facilitate their visual presentation. Table 1 shows the number of late talkers (LT) and typical talkers (TT) in each bin, as well as the mean age in each group.

Stepwise linear mixed effects analysis (*lm4* package, version 1.1-12; Bates, Mächler, Bolker & Walker, 2015) was performed to explore the relationship between each of the three network statistics and type of talker. The final three models includes vocabulary size and type of talker as fixed effects to predict each network statistic. No collinearity of these predictors was found (VIFs < 1.02). Intercepts were allowed to vary across vocabulary sizes as the values of each network statistic varies across the size contexts. The inclusion of vocabulary size as a fixed effect allowed us to control for network size and led to a better statistical fit. No significant difference in the model resulted when allowing slopes to vary by either vocabulary size or type of talker. The significant main effect for type of talker reveals a positive relationship between LT and median indegree (*Estimate*=.061, *SE*=.025, *p*=.014). A marginal main effect for type of talker indicates that being a late talker leads to higher clustering coefficient (*Estimate*=.003, *SE*=.002, *p*=.063). A negative main effect was observed for late talkers on mean distance (*Estimate*= -.014, *SE*=.004, *p*=.0008). These results indicate that LT's vocabularies are better connected and have better global access than TT's vocabularies when considering all the vocabulary sizes together.

We compared the observed networks statistics with those from the vocabulary size-matched random acquisition networks by calculating ratios (Figures 1, 2 and 3). Linear mixed effect analysis was also carried out to examine the relationship between these ratios and type of talker. The model had the same structure as used in the previous analysis.

Results indicate no main effect for the type of talker on in-degree ratio (*Estimate*=.024, *SE*=.014, *p*=.078). Significant main effects were found on clustering coefficient (*Estimate*=.009, *SE*=.005, *p*=.044) and mean distance ratios (*Estimate*=-0.008, *SE*=0.002, *p*=.0007), indicating that late talking children tend to have higher clustering coefficient and lower mean distance in their networks.

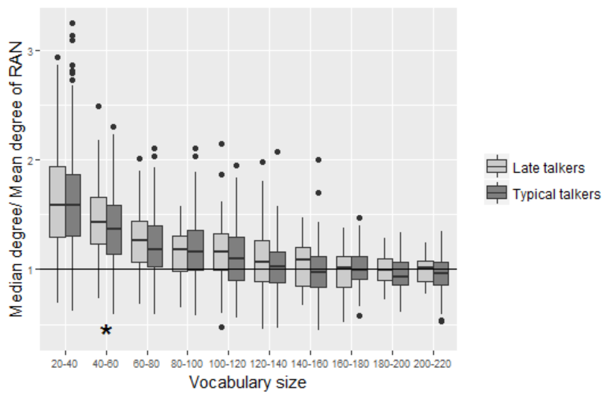

Figure 1. Median degree ratio of the observed data to the RAN. Note: * $p < 0.05$

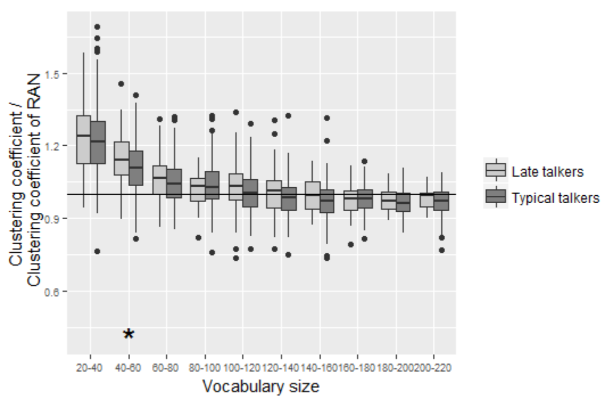

Figure 2. Clustering coefficient ratio of the observed data to the RAN. Note: * $p < 0.05$

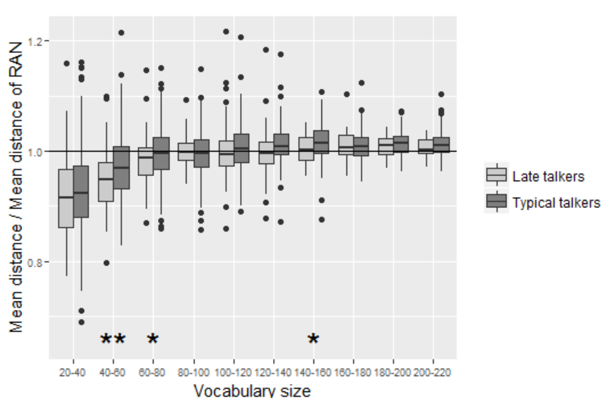

Figure 3. Mean distance ratio of the observed data to the RAN. Note: * $p < 0.05$, ** $p < 0.01$

Table 2: Difference between typical and late talkers compared to random acquisition networks

		Networks M(SD)	RAN M(SD)	t(df)	d
LT	In-degree	23.89 (12.56)	20.81 (12.61)	12.6 (565) ***	.530
	Clustering coefficient	.58 (.07)	.54 (.01)	13.9 (565) ***	.584
	Mean distance	1.68 (.1)	1.73 (.02)	10.33 (565) ***	.434
TT	In-degree	26.7 (13.44)	25.01 (14.56)	13.81(2345) ***	.285
	Clustering coefficient	0.57 (.06)	0.55 (.01)	18.81 (2345) ***	.388
	Mean distance	1.7 (.09)	1.72 (.02)	10.23 (2345) ***	.211

Note: *** = $p < .001$.

Standardized residual plots were visually inspected for the linear mixed models to check for homoscedasticity and normality, and a violation of these requirements was noted. To confirm these results, we carried out more analysis. T-tests were conducted to detect differences between LT and TT regarding the ratio of the observed statistics to the statistics of the size-matched RAN. Late talker's vocabularies have higher values of median in-degree (*M*= 1.28, *SD*= .39) and clustering coefficient (*M*= 1.08, *SD*= .140) than typical talking children (in-degree: *M*= 1.20, *SD*= .37), $t(817) = -4.54$, $p < .001$, $d = -.22$; clustering coefficient: *M*= 1.05, *SD*= 0.13), $t(820) = -4.72$, *p*= $p < .001$, $d = -.23$). Late talking children also had lower values of mean distance (*M*= .97, *SD*= .06) than typical talking children (*M*= .99, *SD*= .063), $t(815) = 5.18$, $p < .001$, $d = .25$.

Further analysis using the ratio data was conducted to check whether the same differences between LT and TT are also observed within each bin of vocabulary size. Significant results are signaled with an asterisk in Figures 1, 2 and 3. Late talkers obtained higher in-degree and higher clustering coefficient than TT in the vocabulary range 40 to 60 words. Significant differences in mean distance were found in three groups of vocabulary size, all of them present lower values for LT: 40 to 60 words, 60 to 80 words, and 140 to 160 words. The same t-test analysis in each bin was performed using only the observed data. Apart from mean indegree, in which any difference between the type of talkers were found, similar results were obtained in clustering coefficient and mean distance in the same vocabulary sizes. These results show that LT's vocabularies are better connected than TT's vocabularies in certain vocabulary ranges, and also that TT resembles their RAN more than LT do in these ranges.

Paired t-tests were conducted to compare the properties of LT and TT networks with their corresponding size-matched RAN. Results can be seen in Table 2. Both LT and TT showed significantly higher values of median in-degree and clustering coefficients and significantly lower values for mean distance than their size-matched RAN. Thus, LT and TT seem to present vocabularies which are well connected and have good global access.

Word Frequency and Connectivity

To investigate differences in the order in which LT and TT learn words, we examined the correlation between word order and word frequency in CHILDES. The order in which all 291 words are learned across vocabulary size is not correlated to word frequency for either type of talker (LT: r_s= -.10, *p*= .082, TT: r_s= -.063, *p*= .28). When the same analysis is performed on each vocabulary size bin, word order in LT was significantly related to word frequency in vocabulary sizes of 20 to 40 words (r_s= -.13, *p*= .023), and 40 to 60 words (r_s= -.16, *p*= .006). Thus, the order in which LT learn words seems to be related to their frequency during the first stages of vocabulary development. Further correlation analysis revealed a relationship between connectivity and word order (r_s= -.13, *p*= .026) in LT but not in TT (r_s= -.11, *p*= .071). Analysis on each bin shows that, only for LT, this relationship is present in vocabulary sizes of 20 to 40 words (r_s= -.16, *p*= .008), 40 to 60 words (r_s= -.18, *p*= .002), 100 to 120 words (r_s= -.13, *p*= .032), 120 to 140 words (r_s= -.13, *p*= .024), 180 to 200 words (r_s= -.12, *p*= .043), and 200 to 220 words (r_s= -.12, *p*= .041). In view of these results, it seems that connectivity and frequency of the 291 words are somehow related to word order in LT but not in TT. In addition, word connectivity seems to be more strongly related to word order in LT than word frequency.

Further analysis considered only those words that differed considerably in word order between LT and TT. That is, there were some words that were learned earlier by one of the type of talker groups compared to the other group, we refer to them here as 'preferred words'. Figure 4 shows significant differences between the two types of talkers: LT learned more words that are highly frequent in the language environment (*M*= 3588, *SD*= 897) than TT (*M*= 1060, *SD*= 687), *t*(16.9) = 7.08, $p < .001$, *d* = 3.16. The connectivity within the matrix of co-occurrences of preferred words was compared between LT and TT. Children with language delay learn words that are well connected in the matrix (*M*= 101, *SD*= 11.4), more so than TT (*M*= 75.8, *SD*= 12.01), *t*(17.9) = 4.79, $p < .001$, *d* = 2.14. These results are consistent with this studies analysis of the network statistics, with LT learning more well-connected words than TT. Results from a logistic regression using the preferred words indicated that frequency is not a good predictor of the type of talker (*B*= .007, *SE*=.008, *p*=.38, *Homes-Lemeshow* R_2=.85), contrary to connectivity, which was a significant predictor (*B*= .14, *SE*= .055, *p*=.009, *Homes-Lemeshow* R_2=.50, 95% CI [1.06, 1.33]).

Additionally, we wanted to see whether word frequency in CHILDES and the connectivity of the 291 words within our matrix of co-occurrences are correlated. Results showed that these two measures are strongly related (r_s= .91, $p < .001$). These results suggest that LT children may be more susceptible to these two word properties, however only the connectivity of the preferred words can predict the type of talker.

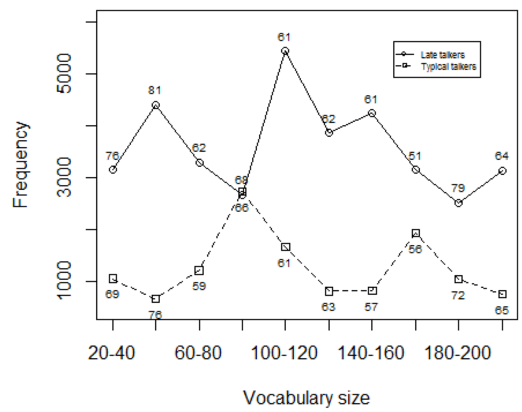

Figure 4. Word frequency by vocabulary size. Numbers at each point reflect number of preferred words used for analysis in each group.

Discussion

The present study uses network analysis on a large sample of children's vocabulary to explore the idea that late talkers may have a different word learning strategy than typical talkers. Multiple results from different analysis fail to agree with the findings in Beckage, Smith and Hills (2011). The authors reported that children with language delay have networks with less clustering coefficient and less mean distance than their vocabulary size-matched RAN. Furthermore, the authors hypothesized that this may be due to LT using an 'oddball strategy' to learn words, i.e. late talkers may be attracted to those words that are not well connected in the learning environment, as opposed to the idea of 'preferential acquisition', in which the tendency is to learn earlier on the most contextually diverse words in the learning environment (Hills, Maouene, Riordan & Smith, 2010). However, the use of a larger sample in this study suggested that LTs do not use an 'oddball strategy', rather they seem to learn words that are well connected in the environment.

Analysis of the frequency of these preferred words in the learning environment showed that LT learn a good proportion of highly frequent words earlier than TT. However, the relationship between word frequency and language delay is still unclear as it seems to not be a good predictor for type of talker. Nevertheless, these results seem to contradict the findings by Stokes (2010), who found that two-year old typical talkers learn more high frequency words than LT do. However, the inclusion of function words may partly explain the difference between the present study's results and Stokes' results as their frequency in the learning environment is higher than open class words. The connectivity of the preferred words within the 291 words used in the present study indicate that LT also happened to learn well connected words earlier.

One of the reasons behind the findings may be that late talking child are more passively influenced by word frequency, consequently, learning more highly frequent words that happen to be well connected in the learning environment. As knowing these high frequency words can deliver a degree of communication success, the requirement to acquire advanced strategies may be further delayed.

In sum, the evidence reported in this study suggests that children with language delay have well connected vocabularies and good global access, in many cases better than the typically developing children. Late talkers may be more influenced by word frequency or connectivity, perhaps using a strategy to learn words that are contextually diverse in the learning environment as noted in previous work (Hills et al., 2009a). In order to elucidate whether these two types of talkers are using different word learning strategies, future research will need to examine the longitudinal development of vocabularies in LT and TT and a different approach to assign a more individualistic semantic relatedness between the words.

References

Banavar, J. R., Maritan, A., & Rinaldo, A. (1999). Size and form in efficient transportation networks. *Nature*, *399*(6732), 130–132.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4, *Journal of Statistical Software*, *67*(1). http://doi.org/10.18637/jss.v067.i01

Beckage, N., Smith, L., & Hills, T. (2011). Small Worlds and Semantic Network Growth in Typical and Late Talkers. *PLoS ONE*, *6*(5). http://doi.org/10.1371/

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition*, *140*, 122–134. http://doi.org/10.1016/j.cognition.2015.03.013

Cancho, R. F. i, & Solé, R. V. (2001). The Small World of Human Language. *Proceedings: Biological Sciences VO - 268*, (1482), 2261

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, *1695*, 1695.

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic Review of the Literature on Characteristics of Late-Talking Toddlers. *International Journal of Language & Communication Disorders*, *43*(4), 361–389. Evert, S. (2008). Corpora and collocations. *Corpus Linguistics. An International Handbook*, 1–53. http://doi.org/10.1515/9783110213881.2.1212

Fenson, L., Dale, P. S., & Reznick, J. S. (1993). *Macarthur communicative development inventories: User's guide and technical manual*. San Diego, CA: Singular Publishing Group.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (in press). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*.

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. http://doi.org/10.1016/j.jml.2010.06.002

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009a). Longitudinal Analysis of Early Semantic Networks Preferential Attachment or Preferential Acquisition? *Psychological Science*, 20, 729-739.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009b). Categorical structure among shared features in networks of early-learned nouns. *Cognition*, *112*(3), 381–396. http://doi.org/10.1016/j.cognition.2009.06.002

Leonard, LB. (2000). *Children with Specific Language Impairment*. Cambridge, MA: MIT Press

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Berlin and New York: Mouton de Gruyter.

Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, *17*(8–9), 1345–1362. http://doi.org/10.1016/j.neunet.2004.07.004

MacWhinney, B. J. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Nematzadeh, A., Fazly, A., & Stevenson, S. (2014). Structural Differences in the Semantic Networks of Simulated Word Learners. *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 1072–1077.

Rescorla, L. (2009). Age 17 language and reading outcomes in Late-Talking Toddlers: Support for a dimensional perspective on language delay. *Journal of Speech, Language and Hearing Research*, 52,16–30.

Stokes, S. F. (2010). Neighbourhood density and word frequency predict vocabulary size in toddlers, *Journal of speech, language and hearing research*, 53, 670–683.

Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in late talkers: A 1-year follow-up. *Journal of Speech and Hearing Research*, 34, 604–612.

Wachs-Lopes, G. A., & Rodrigues, P. S. (2016). Analyzing natural human language from the point of view of dynamic of a complex network. *Expert Systems With Applications*, *45*, 8–22. Retrieved from http://10.0.3.248/j.eswa.2015.09.020

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393, 440-442.