Proceedings of The National Conference
On Undergraduate Research (NCUR) 2018
University of Central Oklahoma
Edmund, Oklahoma
April 5-7, 2018

Phonological and Semantic Features in Foreign Word Learning by Children and Adults

Rebekah Charles
Institute of Child Development
University of Minnesota
Twin Cities, Minnesota

Primary Investigator: Maria Sera, Ph.D.

Abstract

Both phonological and semantic characteristics affect how words are learned in a first language. This study examines these factors in children and adult's foreign word learning. Native English-speaking preschoolers and adults were taught words in Spanish that varied in their phonological and semantic features. They were randomly assigned to one of three semantic conditions and learned words consisting of frequent and infrequent English sound sequences. The words they learned referred to (1) clusters of closely related words such as *lion, rabbit, cat,* and *cow*; or (2) linking words that are associated with two different clusters such as *hat, cookie, ear,* and *cow*; or (3) isolated words that were not related to other words such as *lion, hand, shoe* and *banana*. They were tested on their initial learning of the words and memory for the words about 1 week later. Results indicate that children learned English words consisting of highly frequent sound sequences more effectively than words consisting of infrequent sound sequences. The semantic relatedness of the words did not affect 4-year-olds' performance. However, 6-year-olds were better able to remember words with links to other words better than words from clusters or isolates. Preliminary data from adults suggests that words with low phonotactic probability were easier to remember in session 2, as well as the words in clusters. Findings suggest that different features of language support foreign word learning at different ages.

Keywords: Phonotactic Probability, Semantic Structures, Foreign Language Acquisition

1. Introduction

Understanding how children learn words in a foreign language is important, on theoretical and practical grounds. From a theoretical standpoint, the results are relevant to understanding how a language is learned. Young children are using a number of cues such as joint attention, conceptual knowledge, and native phonology to help them determine the meaning of novel words encountered in their native language. This investigation allows us to explore what aspects of native language learning transfer to the learning of a foreign language.

From a practical standpoint, these results may assist with improving ESL plans and strategies, as there are many students in the educational systems who enter knowing a native language that is not English; According to the National Center for Educational Statistics, 9.4 percent, or an estimated 4.6 million students were ELLs during the 2014-15 school year at public schools in the U.S.[18] Previous research has shown that expanding vocabulary is important in increasing the text comprehension.[3] This study hopes to identify the features of a foreign vocabulary that would facilitate foreign language learning. This could also be relevant to developing more effective curriculum for classes that offer the learning of a second language for native English speakers. It would decrease the amount of time it takes for the student to become proficient as well as how to go about teaching the new words to optimize learning.

Previous studies have investigated factors that influence first language learning, including semantic networks, vocabulary size, object familiarity, and phonotactic probability. Research conducted related to semantic networks suggests that they possess structural features including high sparsity, short average path length and strong local

clustering that should be relevant to word learning.[15] The connections between each node in a semantic network are called paths. Each node is connected to only a small percentage of other nodes, but despite being so sparsely connected, each node is only a small number of path lengths away from the others due to hubs. Hubs are well-connected nodes that link the sparser nodes together. The connectivity patterns lead to strong local clustering, which means that there is neighborhood overlap between the nodes. This clustering and overlap provides evidence of high connectivity of semantic networks, which was theorized by Steyvers and Tenenbaum to facilitate word learning.[15] Supporting evidence indicates that features like neighborhood overlap and strong clustering in children's semantic networks is related to the ability to acquire new words.[4] Typically developing talkers were found to be more likely to acquire words that are semantically associated with words they already know, which supports the claim that learning words is easier when they are semantically related.

However, there have been findings that show that closely related words might have an adverse effect on word learning. For example, Finkbeiner and Nicol found that participants who learned a set of semantically related words in a foreign language learned more slowly and less accurately than the participants who learned a random set of words.[7] This could be attributed to the simultaneous activation of semantically related words, which may cause interference during word-learning tasks. The activation of one node for a particular word could spread to neighboring nodes and impair the ability to differentiate between those semantically dense words within a cluster.

Another factor that has been shown to influence the acquisition of new words in one's native language involves the sounds of the words. Storkel & Lee explored two aspects of this including phonotactic probability and neighborhood density. *Phonotactic probability* (PP) refers to the likelihood that a sequence of phonemes presents appears in a given language. *Neighborhood density* refers to all the words differing by one phoneme substitution, deletion, or any other positioning.[17] For example, neighbors of /sit/ would include words like /hit/, /it/, kin/, and so on. Words with many neighbors would be considered to have high neighborhood density, while those with few neighbors would be considered to have low neighborhood density.[16] Overall, low phonotactic probability and low neighborhood density forms show an advantage in the early mapping of a new word.[9, 17] However, words with high neighborhood density are bettered integrated and retained in the learner's memory.[17]

Storkel and Lee examine the influence of PP on lexical acquisition by 4-year-olds in their native language. Their measures were administered at three time points: prior to training, immediately following the training, and then 1-week after training to examine the retention of the newly created representation. During the training and testing sessions, they continued until either an overall accuracy was met, or the maximum number of trials was completed. The present study differs from that of Storkel and Lee's because this study tested the impact of PP in *foreign* word learning and included adult participants to examine developmental differences.

Research has also been conducted on the effects of phonotactic probability on the performance of wordless repetition (NWR). A robust finding in previous research has concluded that PP, in fact, influences the performance of the NWR. The NWR of high-probability phonological non-words is typically more accurate than NWR of nonwords with low phonotactic probability.[6, 10] There were two reasons discussed, the first was that individuals reconstruct incomplete representations of new stimuli, and because there is more support available for high probability items, repetition is easier. The other explanation was that high probability words are repeated more precisely because phonological processing is favored by more detailed phonological repetitions than those of low probability items.

Sera, Cole, Koening & Oromendia conducted a study looking at the effects of learning familiar versus unfamiliar objects either in their native language, English, or in a foreign language, Spanish.[14] They used a procedure similar to the one used in the present study but they were testing different variables. Three-, four-, and five-year-olds were tested on immediate recollection of the words they had been taught and then they were brought back a week later to test the memory of the words learned. Results show that familiarity with objects facilitates foreign word learning, especially among 3- and 4- year-olds. The present study uses a similar method to examine the effects of phonotactic probability and semantic density on foreign word learning of familiar objects in a foreign language among native English-speaking preschoolers and adults.

2. Methods

2.1 Participants

In this study, native English-speaking 4- and 6-year-olds were recruited from a participant pool at the Institute of Child Development. Families were contacted by phone and were given an explanation of the study, and if the parent/guardian was interested in having their child participate, both sessions were scheduled at that time. An equal number of both

boys and girls without any known mental or physical disabilities, who were native English speakers and were exposed to little or no Spanish, were recruited to participate. There was a total of 48 children who were 4 years old (24 boys, 24 girls), and 48 children who were 6 years old (24 boys, 24 girls). There was also 24 adults (12 male, 12 female) who were recruited and tested to see if further cognitive development influenced the way phonotactic and semantic features affected foreign language acquisition. Most of the adults were students at the University of Minnesota.

2.2 Materials

A box chute was used as a place for the participant to insert the stimuli pictures. A black tray was stationed on the table as a minimally distracting surface to place each pair of pictures on, shown in Figure 1.

Figure 1. Participants picked up the stimuli from the black tray and placed it into the box chute during both sessions.

2.3 Stimuli

Three sources were used to construct the semantic network. The first was The Spoken Word Count (for children ages 5, 6, and 7), collected by Wepman and Hass (1969), which was used to compile a list of nouns that referred to solid objects that would be familiar to the child participants. The second source used was the category-member naming task.[13] Nelson examined recall from long-term semantic memory at ages 5 and 8 by interviewing the children individually and recording their responses when they were given a word from a category and asked to give examples of related words. For example, if given the word from the category *clothes* the responses included were *pants, shirt, dress*, etc. Results showed that 8-year-olds gave twice as many responses as 5-year-olds. The categories with the least number of responses were flowers and vegetables. Nouns for the current study were taken from the responses documented by Nelson.[13] The nouns selected were then cross-referenced with The University of South Florida word association norms to find closely associated words, as well as words that were not related to the targets.[12] The South Florida database consists of 5,019 stimulus words from 6,000 individuals. The adults in the South Florida norming study were asked to report the first word that they thought of that they believed to be closely associated with the stimulus word. The nouns that were semantically related (and likely known by children according to the other two sources) were then used in this study to create the clusters, isolates, and linking items. Four clusters were used (animals, body parts, clothing and food) as shown in Figure 2, each consisting of 4 nouns. Within the clusters, every word had a link to at least two other words within that same cluster; the linking words were closely linked to two words in different clusters. For example, *ear* was related to both *mouth* (in the body parts cluster) and *rabbit* (in the animals cluster. The isolates were only related to one other word within the cluster. For example, *lion* was only related to *cat* within the animal cluster.

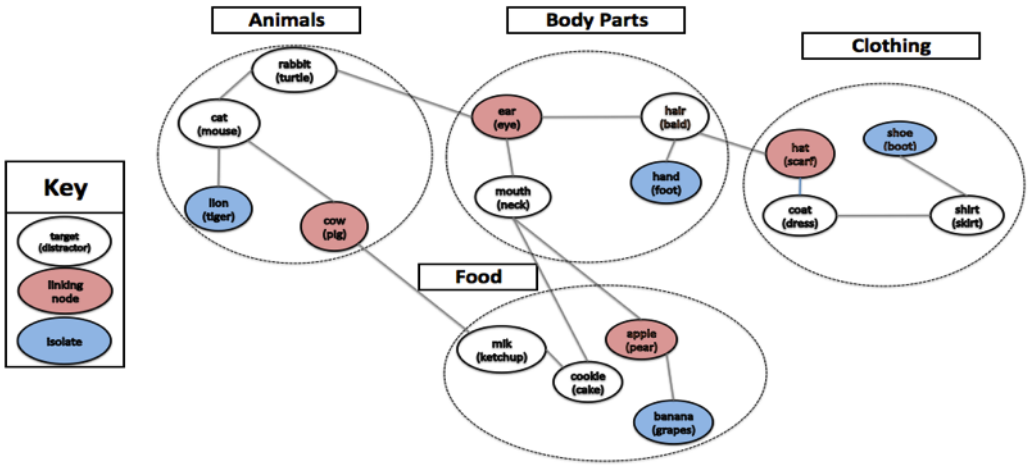


Figure 2. Semantic network, containing 4 clusters comprised of linking nodes and isolates within the target words and distractors.

2.3.1 picture stimuli

Examples of the black and white drawings representing each word are shown in Figure 3. The drawings were copied from the Spanish version of the Peabody Picture Vocabulary Test, Test de Vocabulario en Imagenes Peabody[5] and from the *Pictures, please*![1,2] books contain standardized pictures that are typically used for testing cognitive and language development in children. Distractor words and images were created for each target word, which were pictures of words from the same general networks as the target words, but not as closely associated. Each participant was shown eight pictures per session, four pictures representing the target words, and four representing the distractors. The pictures were presented one at a time in pairs of one distractor and one target picture.

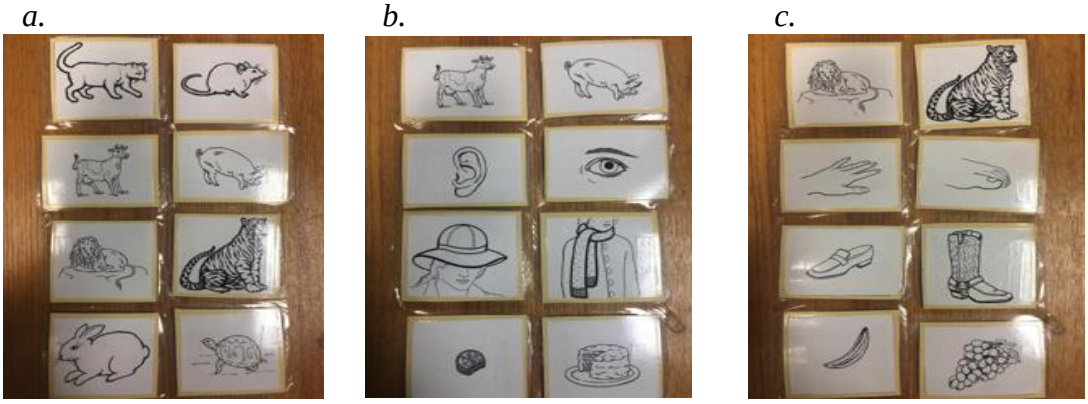


Figure 3. Photo *a* shows the items in the animal cluster, photo *b* shows the items in the linking condition, and photo *c* shows the items in the isolate condition.

2.3.2 nonword stimuli

The nonwords used in this study representing the foreign language (Spanish) were derived from nonwords used in a study by Frisch, Large, and Pisoni.[8] The nonwords were created using sound patterns with either high or low phonotactic probability in English. The original nonword stimuli had two-syllable words with alternating CV patterns, ending in a final consonant. This study took the nonwords created originally for English, and removed the last consonant and replacing the final vowel with either an /o/ or and /a/. Six nonwords were developed, three with

low phonotactic probability and three with high phonotactic probability (see Table 1). The same six nonwords were used in all three semantic conditions.

Table 1. Six nonwords used for all participants, each participant had words of both high and low phonotactic probability.

Phonotactic Probability	
High	**Low**
Mida	Gofa
Henna	Fetcho
Salo	Yuga

3. Procedure and Design

After going over consent forms, the participant was invited into a testing room to go through each session to minimize distractions (when testing children, the child's parent/guardian was invited to enter with child, and asked not to look at any of the stimuli or give any assistance to the child). In each phase, the participant was taught 4 non-words for each of the target pictures. In each condition (cluster, linking or isolate), the tester introduced the target picture and the distractor picture, both of which received the same amount of visual, tangible, and verbal attention (refer back to Figure 2 for examples of the pairs presented). The only difference when presenting the two pictures was that the distractor object did not receive a label. The participants were first given the instructions in English, but the remaining phases of the study were conducted in Spanish only (see the description of Phases below). Only the data of participants who completed both sessions, between 2 and 7 days apart is reported. The first session consisted of three phases, familiarization, training and testing, which are described below. The second session consisted of the testing phase of the items learned in Session 1. The phases were as follows:

3.1 Familiarization Phase

In order to familiarize the participant to the task, the tester explained the process in English. "You'll be playing a game with me, I'll show you two pictures and tell you which one to put inside the chute. We'll go through it once in English, and then I'll switch to Spanish, and you'll do the same thing!" The tester placed two pictures that did not correspond to any cluster onto the black tray, and said "This is a *balloon*, do you see the *balloon*, can you put the *balloon* in here?" while pointing to the picture of the balloon (which would be considered the "target" word) and then to the chute. Once the participant put the correct picture in the chute, the tester pointed to the second picture and said, "Wow, how interesting, do you see it, can you put it in here?" Then the participant was asked, "Where is the *balloon*, do you see the balloon, can you put the balloon in here?" Each statement included the target word three times, and the distractor picture was also mentioned three times without naming the object. After completing the familiarization phase in English, the experimenter switched to Spanish. The experimenter said, "Esto es un *globo.* ¿Vez el *globo?* ¿Puedes poner el *globo* aquí?" ("This is a balloon, do you see the balloon, can you put the balloon in here?") for the target word, and then "Wow, que interesante. ¿Lo vez?, ¿Puedes ponerlo aquí?" ("Wow, how interesting, do you see it, can you put it in here?") for the distractor. Following this, the experimenter asked them to pick out the target word, "¿Donde esta el *globo*? ¿Vez el *globo?* ¿Puedes poner el *globo* aquí?" ("Where is the balloon, do you see the balloon, can you put the balloon in here?").

3.2 Training Phase

After participants were familiar with the task, the tester spoke Spanish for the remainder of the session. The training phase consisted of the same process as the familiarization phase. Two pictures were presented, both the target and distractor (the target picture was placed either on the left or right of the distractor randomly for each of the four pairs). The tester placed the first pair in front of the participant and pointed to the target picture first, Henna for example, saying, "Esto es la *Henna*, ¿vez la *Henna?* ¿puedes poner la *Henna* aquí?"("This is the Henna, do you see the Henna, can you put the Henna here?") After the participant put the Henna in the chute the tester then pointed to the distractor

saying, “Wow, mira eso, que interesante, ¿puedes ponerlo aquí?”(“Wow, look at that, how interesting, can you put it in here?”). This is done with the remaining three pairs, after which concludes the training phase.

3.3 Testing Phase

The testing phase took place in Spanish, using the same nonwords for the target pictures as in the training phase. The tester presented each pair again, this time asking the participant to identify the target picture, i.e. “Cual es la *Henna*, ¿puedes ver la *Henna?* ¿Puedes poner la *Henna* aquí?” (“Which one is the henna, can you see the *Henna*, can you put the *Henna* in here?”). If the participant put the correct picture in the chute the tester would move on to the next pair, however, if the participant put the distractor picture in the chute, the tester would say “No, esto es la *Henna*, ¿puedes ponerlo aquí?” (“No, this is the *Henna*, can you put it in here?”) and would then go on to the next pair. Going through all four pairs was considered the immediate test trial, and if the participant got one wrong, the four trials would be repeated up to four times or until the participant got all four target pictures correct. The order of the target picture and distractor picture was randomized during both the training and testing phase.

4. Results

The percentage of words correctly learned and remembered of High versus Low PP were compared in all three age groups for both sessions and across all three semantic conditions through ANOVAs. The findings for each age group are reported below.

4.1 Four-year-olds

High PP helped 4-year-olds to immediately learn the words in Session 1 ($p=.034$). No reliable effects of semantic structure were found for 4-year-olds in Session 1. There were no reliable effects of semantic or phonological structure on the ability of 4-year-olds to remember the words in Session 2.

4.2 Six-year-olds

High PP also helped the 6-year-olds to immediately learn the words in Session 1 ($p = .023$). No evidence was found showing that the semantic structure influenced their ability to initially learn the words. In contrast to 4-year-olds, however, 6-year-olds benefitted from High PP in remembering the words in Session 2 ($p=.017$). Six-year-olds’ memory for the words that were linked to two different clusters was better than their memory for the words in clusters or isolates ($p = .0015$).

4.3 Adults

Adults were able to easily learn all the words across all of the conditions in Session 1 and showed no differential sensitivity to the phonological or semantic structure in initial learning. However, the evidence from adults suggests that words with low PP were easier to remember for them in Session 2, as were the words in clusters. However, power analyses suggested that the adult sample size was too small to yield significant results.

Table 2. Percent of words correctly recalled by participants in Session 1.

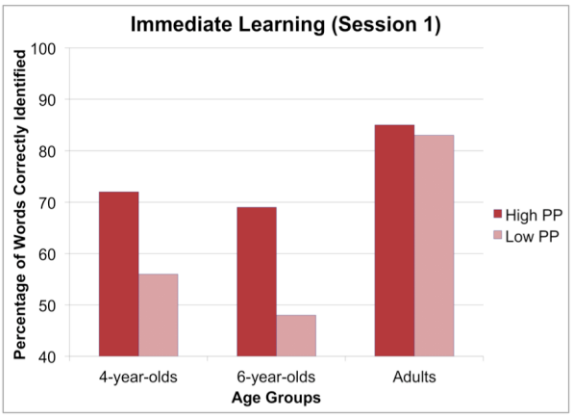
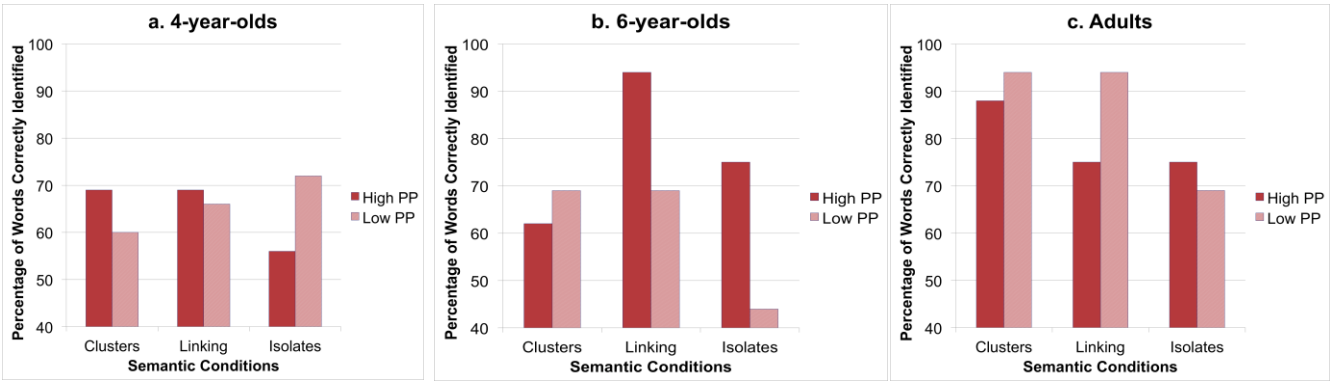


Table 3. Percent of words correctly recalled by (a) 4-year-olds, (b) 6-year-olds and (c) adults in Session 2 (memory) across all conditions.



5. Discussion

This research began by asking what effects phonotactic probability (PP) and semantic structure (SS) had on young children's foreign word learning. By examining how 4-year-olds, 6-year-olds, and adults learn words for familiar objects in a foreign language, we were able to shed light on what aspects of a native language learning transfer to the learning of a second language across development. Previous research has looked at these same characteristics but has solely focused on the role of these factors in first language learning. This study examined the role of these factors in native English speaking children and adult's ability to learn words in a foreign language.

Results suggest that in children, immediate foreign word learning is linked to High PP, indicating that children's initial word learning is influenced by the degree to which the foreign words match the sound structure of the words in their native language. The four-year-olds' semantic networks may not be sufficiently developed to benefit (or be hindered) from semantic relatedness. They may not have the links between the items in their networks; or by having fewer items in their networks, their items might all be so close together that they interfere with each other.

Six-year-olds also benefitted from learning foreign words that matched the sound structure of the words in their native language, and these effects of sound structure carried into their ability to remember the words. Results showed that by the age of 6 the semantic structure affects the ability to remember foreign words. Interestingly, 6-year-olds did not benefit most from the words in clusters that were most similar to each other; perhaps the large similarity between

these words interfered with each other in memory. Instead, they were better able to remember a foreign word if it was semantically linked to other items in their network, words that were related to non-similar items (such as *ear* and *rabbit*). These findings from 6-year-olds contrast to the preliminary findings from adults whose memory for the words benefitted most for items that were most closely related (the clusters).

Previous studies have found mixed effects when looking at word characteristics such as PP or SS in a single language, showing that isolated things with lower PP are easier to remember.[11, 17, 19] The difference between previous studies and our current findings are that our results are based on the learning of a *foreign* language. Monolingual studies supporting the facilitation of word learning through low PP or low SS may be occur because if the study is conducted in your L1 (language one), it would seem hearing low PP words would grab your attention, whereas in a foreign language study, such as this one, everything is conducted in L2, which seems to increase the ability to pick out the high PP sounds that are most familiar. Our findings also indicate that different aspects of the structure of language benefit individuals of different ages. The youngest children only benefitted from the similarity in sound structure; older children also benefitted from the sound structure but also began to benefit from semantic structure. The semantic structure also benefitted adults, but a larger degree of similarity was needed for them to reap a benefit.

These findings are also relevant to practical questions regarding how and when to introduce children to a foreign language. Our findings suggest that teaching children high PP words would optimize foreign word learning. As children develop, introducing them to words with greater links between them may yield positive results. This can be applicable to children with English as their L1, learning a foreign language, but also to English Language Learners by creating more effective curriculum and more strategic programs. Because ELL curriculum tends to run from K-12, future studies may want to repeat this study with Elementary, Middle, and High School age groups to see if there are significant differences between other age groups. This may also include looking at the relationship between previous and new learning, as well as why certain semantic conditions facilitated foreign word learning among the older age groups.

In conclusion, our results suggest how to support foreign word learning in children. Both age groups benefited from words that consisted of frequent native sound combinations. Teaching words to children with these sound characteristics should enhance their ability to learn words in a foreign language, which includes English for those who speak another language at home. The semantic relationships between the words did not affect the ability of younger children to learn foreign words. Older children, however, benefited from learning words that were associated with other words. The preliminary findings from the adult data suggest that teaching them words with close semantic relationships and infrequent sound sequences will facilitate their learning, however, more research is needed to support these results. In short, our findings illustrate how different linguistic characteristics can support foreign language learning at different ages.

6. References

1. Abbate, M. S., & La Chappelle, N. B. (1984a). Pictures, please! A language supplement. Tucson, AZ: Communication Skill Builders.
2. Abbate, M. S., & La Chappelle, N. B. (1984b). Pictures, please! An articulation supplement. Tucson, AZ: Communication Skill Builders.
3. August, D. et al. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50-57.
4. Beckage, N., Smith, L., Hills, T. (2011). Small world and semantic network growth in typical and late talkers. *PLUS ONE* 6(5): e19348. doi:10.1371/ journal.pone.0019348.
5. Dunn, L. M., Padilla, E. R., Lugo, D. E. & Dunn, L. M. (1986). Examiner's manual for the test de vocabulario en imágenes peabody: Adaptación hispanoamericana. Circle Pines, MN: American Guidance Service, Inc.
6. Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*, 421-436.
7. Finkbeiner, M., & Nicol, J. (2003). Semantic category effects in second language word learning. *Applied Psycholinguistics*, 24, 369-383.
8. Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language,* 42(4), 481-496.

9. Hoover, J. R., Storkel, H. L., & Hogan, T. P. (2010). A cross-sectional comparison of the effects of phonotactic probability and neighborhood density on word learning by preschool children. *Journal of Memory and Language*, *63*(1), 100-116.
10. Messer, M. H., Leseman, P. P. M., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relation with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology, 105,* 306-323.
11. McKean, C., Letts, C., & Howard, D. (2014). Triggering word learning in children with language impairment: The effect of phonotactic probability and neighbourhood density. *Journal of Child Language*, *41*, 1224–1248.
12. Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.
13. Nelson, K. (1974). Variations in children's concepts by age and category. *Child Development*, *45*(3), 577-584.
14. Sera, M. D., Cole, C. A., Koenig, M. A., & Oromendia, M. (2014). Object familiarity facilities foreign word learning in preschoolers. *Language Learning and Development. 10*(2), 129-148.
15. 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.
16. Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics, 25*(2), 201-221.
17. Storkel, H. L., & Lee, S. Y. (2011). The independent effects of phonotactic probability and neighborhood density on lexical acquisition by preschool children. *Language and Cognitive Processes, 26*(2), 191-211.
18. U.S. Department of Education, National Center for Education Statistics. (2017). *The Condition of Education 2017* (2017-144), English Language Learners in Public Schools.
19. Van der Kleij, S. W., Rispens, J. E., Scheper, A. R. (2016). The effect of phonotactic probability and neighbourhood density on pseudoword learning in 6- and 7-year-old children. *First Language, 36*(2), 93-108.