A computational model of three non-word repetition tests

Gary Jones (Gary.Jones@NTU.ac.uk)
Division of Psychology, Nottingham Trent University, Chaucer Building, Goldsmith Street
Nottingham, NG1 5LT (UK)

Abstract

The non-word repetition test has been regularly used to examine children's vocabulary acquisition, and yet there is no clear explanation of all of the effects seen in non-word repetition. This paper presents a study of 25 5-6 year-old children's repetition performance on three non-word repetition tests that vary in the degree of their lexicality. EPAM-VOC, a model of children's vocabulary acquisition, is then presented that captures the children's performance in all three repetition tests. The model represents a clear explanation of how working memory and long-term linguistic knowledge interact in a way that is able to simulate performance in non-word repetition.

Keywords: Computational modelling; Non-word repetition; Child development.

Introduction

One ability that sets the human species apart from other species is that of language. However, the learning of language is a complicated process that involves at least the following processes. First, the learner must identify where words begin and end from speech that is often continuous. Second, the learner must store the newly identified words in their long-term lexicon. Finally, the learner must acquire the rules of syntax and grammar that govern the way in which their lexicon words can be combined. It is the second of these three processes that this paper is focused: the process of vocabulary learning.

Research that examines vocabulary learning is proliferated with tests of non-word repetition – a test that involves nonsense words being spoken aloud to the language learner, who must repeat them accurately. The test involves non-words since one can be certain that the child has never encountered the sequence of sounds before, hence providing a true test of vocabulary learning. Furthermore, studies of vocabulary involving non-word repetition have primarily focused on children, since the vast majority of language learning occurs early in one's development.

Non-word repetition research

Non-word repetition tests were originally developed to examine the influence of phonological working memory on the vocabulary learning process. For example, Gathercole and Baddeley (1989) showed that repetition accuracy improved between the ages of 4 and 5 years, and performance declined as non-word length increased for both ages. Both of these findings were interpreted in terms of phonological working memory: an improvement with age could be explained by an increase in memory capacity; and a decrease in performance as non-word length increased could be explained by the decay of items in working memory.

However, subsequent research has shown that the child's existing lexical knowledge plays a major role in their non-word repetition ability. Gathercole (1995) re-analysed the non-words in the original test by separating them into "wordlike" and "non-wordlike" non-words based on adult subjective ratings of wordlikeness. She found that children performed significantly better for non-words that were wordlike. Although wordlikeness is a subjective measure, even when more objective measures are used, there are still clear differences between non-words that share substantial lexical features with words compared to those that do not. For example, if one actively distinguishes non-words based on their constituent phoneme combinations – having one set that contain highly frequent combinations of sounds versus a set containing relatively infrequent combinations – there are clear performance differences, with children regularly finding the high-frequency non-words easier to repeat (e.g. Edwards, Beckman & Munson, 2004; Vitevich, Luce, Charles-Luce & Kemmerer, 1997).

It would therefore seem that non-word repetition, and in turn vocabulary acquisition, can be affected by both phonological working memory and long-term lexical knowledge. There are at least two prominent explanations of vocabulary acquisition that explain repetition performance in terms of both processes.

Explanations of non-word repetition performance

Since non-word repetition performance is affected by an interaction between working memory and long-term memory, any explanation of performance must provide some detail of how these two processes interact. Gathercole (2006) explained this interaction using the idea of phonological frames. Phonological working memory is used to store linguistic stimuli (e.g. non-words in the repetition test) and when these items decay, long-term linguistic knowledge is relied upon to help bolster the decaying representations in working memory. Since non-words that are wordlike, or that contain highly-frequent sounds, will share more information with lexical items in long-term memory, it is these items that gain more help from existing vocabulary knowledge. That is, the support provided by the phonological frames of existing vocabulary items increases as the amount of overlap in shared features (to non-words) increases.

An alternative explanation of vocabulary learning shares many features with Gathercole's idea of phonological frames, yet is more explicit in its detail. Jones, Gobet and Pine's (2007, 2008) EPAM-VOC is a computational model

of vocabulary learning that concretely specifies how phonological working memory and long-term phonological knowledge interact. Long-term knowledge is represented by "chunks" of phoneme sequences – as the model is subjected to more and more linguistic input, these chunks of phonemes become larger and larger. Phonological working memory is represented by a fixed amount of chunks that can be stored. Hence, early on in the model's learning, EPAM-VOC is able to store only a limited amount of linguistic information in working memory since the chunks at this point in time will not be large sequences of phonological information. Later on in learning, the phoneme sequences within chunks will be relatively large, and so an increased amount of information can be stored in working memory even when the number of chunks remain the same. This explanation of vocabulary learning puts forward the idea that improved performance with age arises due to an increased amount of linguistic knowledge. However, the model also explains wordlikeness and frequency effects quite easily. Phoneme sequences that appear regularly in the child's language will be represented within the model as relatively large chunks, whereas low frequency sequences will not. Therefore non-words that contain high frequency sequences can be stored in working memory using few chunks, giving rise to an increase in the likelihood of their correct repetition over non-words containing low frequency sequences. A similar explanation can be used for wordlike versus non-wordlike non-words. The former, since they bear great resemblance to words, are likely to be represented within the model using fewer chunks than non-wordlike non-words.

The current paper

EPAM-VOC has thus far been used to successfully simulate the non-word results of Gathercole and Baddeley (1989) plus a non-word set devised for children between 2 and 5 year of age. However, neither of these tests were specifically designed to vary in their lexicality. Since research has shown that non-word repetition is strongly influenced by the lexical nature of the non-words involved, this paper examines EPAM-VOC's repetition performance across three sets of non-words that vary in the degree of their lexicality. The remainder of this paper is as follows. First, the model itself is detailed so that the reader has more extensive knowledge of its workings. Second, a new study of 5-6 year-old children's repetition is described that presents three different non-word repetition tests that vary in the extent of their lexicality. Third, the results of the children are compared to the results of the model. Finally, a discussion of the results are given.

EPAM-VOC: A model of vocabulary learning

EPAM-VOC is a model of vocabulary learning that is based on the EPAM modelling architecture (Feigenbaum & Simon, 1984). This architecture has been used to successfully simulate human behaviour in a range of psychological domains (see Gobet et al., 2001).

Furthermore, the modelling environment has been successfully applied to syntax acquisition as well as vocabulary acquisition (e.g. Freudenthal, Pine & Gobet, 2006; Freudenthal, Pine, Aguado-Orea & Gobet, 2007). The model presented here is an updated version of that described by Jones, Gobet and Pine (2007, 2008), since that model did not have an explicit articulation process.

Representing long-term knowledge

Knowledge within EPAM-VOC is represented as a hierarchy of chunks that contain phoneme sequences. Chunks that are lower down in the hierarchy contain larger sequences, and hence EPAM-VOC can be equated to a tree-like structure. The model begins with knowledge of all of the constituent phonemes in English, since there is good reason to believe that even at an early age, children have knowledge of the phonemes of their language (Bailey & Plunkett, 2002).

An example hierarchy of chunks is given in Figure 1. Here it can be seen that the model knows the phoneme sequence for the word "Toys" (T OY Z). Note that we represent phonemes using the phoneme set in the CMU Lexicon Database (available at www.speech.cs.cmu.edu/cgi-bin/cmudict) rather than the International Phonetic Alphabet. This is chiefly because the database allows the semi-automatic conversion of spoken utterances into their phonemic equivalent (this will be detailed later when the input regime for the model is covered).
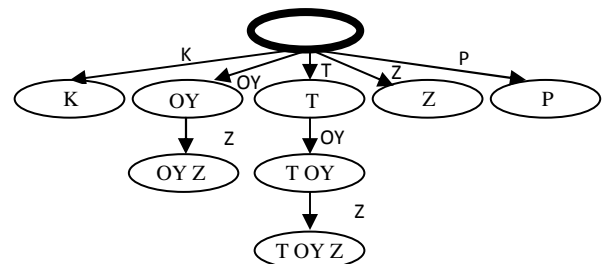

Figure 1: Graphical representation of EPAM-VOC having been presented with "Toys" (T OY Z) twice as input. Chunks are represented by ellipses and links are represented by arrows. Note that although only five phonemes are represented as single phoneme chunks (K, OY, T, Z and P) the model knows all phonemes in English as individual chunks.

Representing phonological working memory

Working memory in this version of EPAM-VOC is no longer limited to a set amount of chunks. Instead, there is a set amount of activation that is spread over the chunks that are in working memory (c.f. Cowan, 1997). However, we base this activation on time, since there is solid research to indicate that items in working memory have a temporal duration of 2,000 ms unless rehearsed (e.g. Baddeley, Thomson & Buchanan, 1975) and there is solid research that

places timing estimates on accessing a chunk and accessing its constituent phonemes (Zhang & Simon, 1985).

For any given input, EPAM-VOC's long-term knowledge is accessed in order to convert the input into a series of chunks (i.e. representing the input sequence in as few chunks as possible). Each chunk is then assigned an access time – 400 ms to access the chunk and a further 30 ms to access each constituent phoneme bar the first (e.g. given the long-term knowledge of Figure 1, "Toys" would be allocated 460 ms, whereas the single phoneme "K" would be allocated 400 ms). Once the input has been represented in as few chunks as possible, and each chunk has been assigned an access time, then a pointer to each chunk is placed in working memory. The total access time is calculated by summing the access times for all chunks. When this total exceeds 2,000 ms, then there is a probability of less than 1.0 that a chunk can be correctly accessed from its pointer (when learning or articulating, the model requires the chunk to be accessed from its pointer in working memory).

Let us take the input "My toys are here" as an example (phonemic representation: "M AY T OY Z AE R H IY R") and the knowledge in Figure 1. Only "T OY Z" exists as a multi-phoneme chunk, and this is assigned an access time of 460 ms. All other phonemes ("M", "AY", "AE", etc.) are assigned an access time of 400 ms – there are a total of 8 chunks required to represent the input, in a total access time of (7*400 ms)+(1*460 ms)=2,560 ms. The probability of subsequently accessing a chunk from its pointer is the temporal duration of working memory divided by the total time required to access all of the chunks: 2,000/2,560=.78125.

To summarise, any given input is converted into as few chunks as possible using EPAM-VOC's long-term knowledge of phoneme sequences. This matching process takes a certain amount of time, and the result of the process is that a pointer to each chunk is placed in working memory. Since working memory only contains pointers to chunks, any process that requires the actual information in the chunk (e.g. when learning or articulating items in working memory) must access the chunk itself. The accurate access of information in a chunk is probabilistic, dependent upon whether the total access time for all chunks exceeds the 2,000 ms time limit of working memory.

Learning phoneme sequences

During learning, any given input is coded into as few chunks as possible, and pointers to the chunks are placed in working memory (as described above). The learning process then examines each pair of pointers to see if a phoneme sequence can be learnt that combines the information within each chunk pairing. This can only be done if each chunk is correctly accessed, but if this is the case, a new chunk is learnt whose contents are the combination of both chunks. Let us use the input "Toys" ("T OY Z") as an example. When EPAM-VOC first begins its learning, it only knows single phonemes as chunks, and therefore "T OY Z" would

be represented in working memory using three pointers to three chunks (one pointer to each of "T", "OY" and "Z"). Since the time to encode the three chunks is 400 ms for each (totalling 1,200 ms and therefore within the 2,000 ms limit) then the subsequent accessing of the information within the chunks will be completely accurate. EPAM-VOC takes each pair of pointers in turn and tries to learn something from them. The first pair are "T" and "OY". The "T" chunk is accessed, and then a link to a new chunk is placed below the "T" chunk. The link will specify the additional information that is being learnt ("OY") and the new chunk contains the joint set of information ("T OY"). The next pair of chunks ("OY" and "Z") are then examined, and in a similar vein, a new chunk "OY Z" is learnt. Should "T OY Z" be presented to EPAM-VOC a second time, it can now be represented as two pointers to the chunks "T OY" and "Z". The learning from this pair of pointers would result in a new chunk "T OY Z" being added below the "T OY" chunk, and the resulting network would be that shown in Figure 1.

Let us now see how learning progresses when the access time exceeds the 2,000 ms limit. Take the previous example sequence "My toys are here" ("M AY T OY Z AE R H IY R") and the long-term knowledge of Figure 1. It was already determined that there was a .78125 probability of accessing a chunk that related to a pointer for this input. Since the pointers in working memory are analysed in a pairwise fashion, then if one pointer cannot access its associated chunk, no learning can be accomplished for that pointer. For example, if the pointer to the chunk "AY" failed, then EPAM-VOC could not learn the sequence "M AY" or the sequence "AY T OY Z".

Articulating phoneme sequences

In order for a phoneme sequence to be articulated, it must first be represented in working memory as a series of pointers to chunks (as described above). In the same way as for learning, each chunk needs to be correctly accessed from its pointer, otherwise an incorrect articulation takes place. However, even if each chunk is correctly accessed, the chunk may still be incorrectly articulated based on its frequency. EPAM-VOC maintains a frequency for each chunk based on the number of times that the chunk has been accessed. We assume that the articulation of phonemes in a chunk is based on the frequency of that chunk – those chunks that are low in frequency will attract more errors than chunks that are high frequency. Correct articulation of an input sequence (e.g. a non-word) is only achieved when all of the relevant chunks are correctly encoded into phonological working memory, and all of the phonemes are correctly articulated from each chunk based on the frequency of the chunk.

Training the model

The model uses naturalistic speech input based on the maternal input from 12 sets of mother-child interactions to 2-3 year-old children (Theakston, Lieven, Pine & Rowland, 2001). All input is converted into the phonetic alphabet of

the CMU Lexicon Database, as discussed previously. 12 simulations are carried out, one for each set of mother's input. However, since comparisons are going to be made to 5-6 year-old children, additional input was sought from paternal interactions with 5 year-old children plus input from reading material for children of this age group (e.g. Snow White).

During training, the model was presented with the same amount of input as per the original maternal input. However, as learning progressed, more and more of the maternal input was replaced with input that reflected that which a 5-6 year-old child would receive.

Since the input to the model can vary based on which utterances from the mother were chosen for replacement, and which input from the 5-6 year-old input set was chosen as the replacement, then the model was run ten times for each "mother". This ensures that the results from the model are replicable and are not simply based on an advantageous input set. Similarly, the non-word repetition tests are carried out ten times for each simulation, since there are probabilistic elements to both encoding and articulation. There were therefore, for any non-word in a non-word repetition test: 12 mothers * 10 simulations runs * 10 attempts at each non-word = 1,200 repetitions of each non-word.

5-6 year-old children's repetition performance

Participants
25 5-6 year-old children (5;4-6;8, M=6;1; 10 male, 15 female) who all scored within normal ranges on the British Picture Vocabulary Scale (Dunn, Dunn, Whetton & Burley, 1997). All children were English monolinguals and had no hearing difficulties, as reported by their school teacher.

Materials
The CNRep (Gathercole, Willis, Baddeley & Emslie, 1994) that includes non-words that are considered high in lexicality since they include syllables that are either real words (e.g. **thick**ery) or morphemes (e.g. glister**ing**). The non-words in this test are either single consonant (e.g. sladding) or clustered consonant (e.g. glistow). There were 5 non-words of each type at each of three lengths (2-4 syllables). 5-syllable non-words were excluded because children at this age had difficulty in repeating non-words of this length.

The non-word test of Dollaghan, Biber and Campbell (1995) that contain 3-syllable non-words. 6 non-words contained a real-word syllable (e.g. **bath**esis) and 6 changed one phoneme in the non-word so that it was entirely non-lexical (e.g. fathesis). This test is considered medium in lexicality.

A new non-word repetition test that was entirely non-lexical and that contained two sets of 3-syllable non-words: 8 that were low frequency (e.g. latmunoz) and 8 that were very low frequency (e.g. wegnerterk). The non-words in each set were matched for syllable structure, number of phonemes, age of acquisition of the phonemes, and number of consonant clusters. The average log frequency was lower for very low frequency non-words than low frequency ones (.51 vs. .44, $t(7)=3.92$, $p<.01$) using a procedure for measuring bi-phone frequency similar to that of Luce and colleagues (e.g. Jusczyk, Luce & Charles-Luce, 1994; Vitevich, Luce, Charles-Luce & Kemmerer, 1997). This test was considered to be low in lexicality.

Design
The CNRep had two independent variables: non-word type (single or clustered) and non-word length (2, 3, or 4 syllables). The Dollaghan non-words had one independent variable (lexicality: lexical or non-lexical). The new non-word test also had one independent variable (frequency: low or very low). The dependent variable in all cases was repetition accuracy.

Procedure
All children were tested individually on a one-to-one basis in a quiet area of their school. Each non-word repetition test was carried out on a separate day. For the CNRep, the original recordings were maintained, but for the other two repetition tests the non-words were recorded by a speaker native to Nottingham. The instructions for each set of non-words are given below, and were the same for each non-word test. Children's responses were recorded onto a Sony ICD-MX20 digital voice dictaphone for later analysis.
"Hello, in a few seconds you will hear a funny made up word. Please say the word aloud yourself as soon as you hear it."

Results
For each repetition test, two sets of results are shown for the model: the average of all of the 1,200 simulations, plus the average of 12 simulation runs (one from each mother). The 12 runs are included since statistical analyses are based on these – the nature of the 1,200 simulations means that they show little variance, since they are all based on a similar set of input data. The selection of the single simulation on which to base statistical analyses was pseudo-random – that is, the individual 1,200 simulation runs were narrowed down to those that approximated the average of all 1,200 runs when taken as a whole, and one run (the first run for each mother together with the seventh of the ten duplicate repetitions) was randomly chosen from that set.

Figure 2 shows the children's results for the CNRep together with the results from EPAM-VOC. A 2 (non-word-type: single or clustered) x 3 (non-word-length: 2, 3, or 4 syllables) repeated measures ANOVA was performed on the children's data. There was a significant effect of non-word-type ($F(1,24)=43.5$, $p<.001$), showing that performance was better for single consonant non-words, and a significant effect of non-word-length ($F(2,24)=26.7$, $p<.001$), showing that performance was better for short non-words. There was no interaction between the two ($F(2,48)=1.8$, $p>.05$). For the

model, there was also the same effect of non-word-type ($F(1,11)=5.5$, $p<.05$) and non-word length ($F(2,22)=9.8$, $p<.001$), with no interaction between the two ($F(2,22)=.3$, $p>.05$).
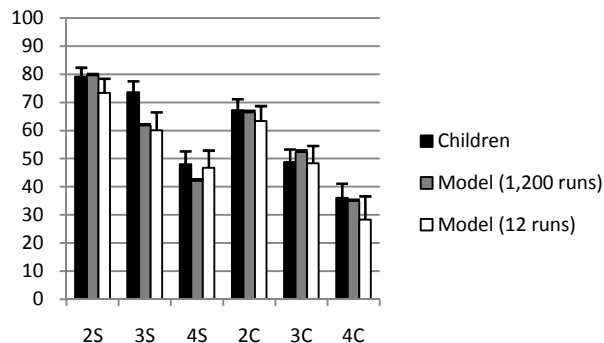

Figure 2: Non-word repetition performance (%) for the CNRep and for the two sets of model runs. The numeric on the x-axis denotes the number of syllables (2, 3, or 4) and the alphabetic character denotes the non-word type (S=single consonant, C=clustered consonant).

Figure 3 shows the children's and model's performance for the Dollaghan non-words and the new set of non-words. For the Dollaghan non-words, the children showed no difference in their ability to repeat lexical and non-lexical non-words ($t(24)=.6$, $p>.05$). The same was found in the model ($t(11)=.6$, $p>.05$). For the new set of non-words, there was no difference in children's repetition accuracy between low and very low frequency non-words ($t(24)=.1$, $p>.05$). Again, the same result was found in the model ($t(11)=.5$, $p>.05$).
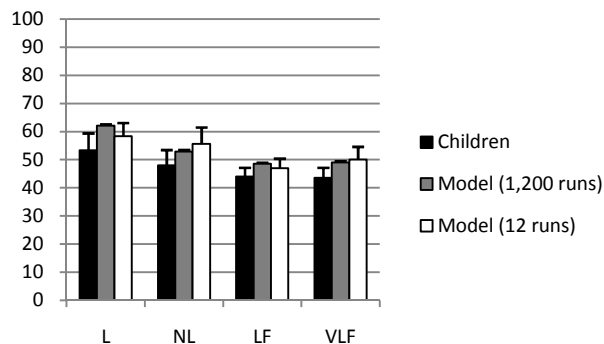

Figure 3: Non-word repetition performance (%) for the Dollaghan and new non-words, for the children and the two sets of model runs. L=Dollaghan, Lexical non-words; NL=Dollaghan, Non-Lexical non-words; LF=Low Frequency new non-words and VLF=Very Low Frequency new non-words.

Discussion

Figures 2 and 3 show that the model provides a very close fit to the repetition data of 5-6 year-old children. The central finding is that the statistical analysis of the model's data mirrors that of the children: clear effects are found for the non-words of the CNRep whereas no effects are found for the other repetition tests. The results from each set of non-words will now be discussed in turn.

The CNRep results are exactly those found in 4 and 5 year old children (e.g. Gathercole & Baddeley, 1989): performance improves for single consonant non-words and for non-words of fewer syllables. In fact, this set of findings is rather robust since they persist in older age groups also (e.g. Briscoe, Bishop & Norbury, 2001). Both the previous version of EPAM-VOC and the new version presented here are able to simulate these findings, suggesting that a reasonable account of working memory and its interaction with long-term linguistic knowledge is sufficient to capture the behaviour shown in the CNRep.

For the Dollaghan non-words, the original study showed an effect of lexicality for 10 year-old children (Dollaghan, Biber & Campbell, 1995). Not only do the children in this study not show this lexicality effect, the model itself also does not capture it. The model puts forward an explanation for the lack of effect, in that the lexical items (e.g. bath) are not robust enough in terms of their frequency of use to cause improved performance for non-words containing a lexical item. It would be interesting to take the learning in the model a stage further to the type of input older children may receive to then see if lexical effects emerge.

For the new set of non-words, there was no effect of frequency in either the children or the model. This shows that frequency effects are not picked up by children of this age, although they may well be for older children.

In summary, EPAM-VOC replicates the findings of 5-6 year-old children on three different non-word repetition tests varying in the degree of their lexicality. It now needs to be seen whether the errors made in children's repetitions are also mirrored by the model – if this is the case, then EPAM-VOC may prove to be a very strong explanation of the way in which children are learning vocabulary.

Acknowledgments

References

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour, 14*, 575-589.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17*, 1265-1282.

Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, 42*, 329-340.

Cowan, N. (1997). The development of working memory. In N. Cowan & C. Hulme (Eds.), *The development of memory in childhood* (pp. 163-199). Hove, UK: Psychology Press.

Dollaghan, C. A., Biber, M. E., & Campbell, T. F. (1995). Lexical influences on non-word repetition. *Applied Psycholinguistics, 16*, 211-222.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale (2ⁿᵈ ed.).* Windsor, UK:NFER-Nelson.

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 non-word repetition. *Journal of Speech, Language, and Hearing Research, 47*, 421-436.

Feigenbaum, E. A. & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science, 8,* 305-336.

Freudenthal, D., Pine, J.M., Aguado-Orea, J. & Gobet, F. (2007). Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science, 31*, 311-341.

Freudenthal, D., Pine, J.M. & Gobet, F. (2006). Modelling the development of children's use of Optional Infinitives in English and Dutch using MOSAIC. *Cognitive Science*, *30*, 277-310.

Gathercole, S. E. (2006). Non-word repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.

Gathercole, S. E. (1995). Is non-word repetition a test of phonological memory or long-term knowledge? It all depends on the non-words. *Memory & Cognition, 23*, 83-94.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of non-word repetition: A test of phonological working memory. *Memory, 2*, 103-127.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200-213.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C-H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *TRENDS in Cognitive Sciences, 5*, 236-243.

Jones, G., Gobet, F., & Pine, J. M. (2008). Computer simulations of developmental change: The contributions of working memory capacity and long-term knowledge. *Cognitive Science*.

Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of novel sound patterns. *Developmental Science, 10*, 853-873.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*, 630-645.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language, 28*, 127-152.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech, 40*, 47-62.

Zhang, G. & Simon, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. *Memory & Cognition, 13*, 193-201.