

A computational simulation of children's performance across three nonword repetition tests

Gary Jones

Division of Psychology, Nottingham Trent University, Chaucer Building, Goldsmith Street

Nottingham, NG1 5LT (UK)

Gary.Jones@NTU.ac.uk

Abstract

The nonword 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 nonword repetition. This paper presents a study of 5-6 year-old children's repetition performance on three nonword 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 lexical and sub-lexical knowledge interact in a way that is able to simulate repetition performance across three nonword tests within the same model and without the need for test specific parameter settings.

Keywords: Computational modelling; Nonword repetition; Phonological working memory; EPAM-VOC.

Introduction

One ability that sets the human species apart from other species is that of language. However, the learning of language is a complex procedure 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 rules of syntax 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 nonword repetition – a test that involves nonsense words being spoken aloud to the language learner, who must repeat them accurately. Nonwords are used since one can be certain that the child has never encountered the sequence of sounds before, hence providing a true test of vocabulary learning. Studies of nonword repetition have found clear links between repetition performance and vocabulary learning. For example, Gathercole and Baddeley (1989) measured repetition performance at four years of age and found it to be the best predictor of vocabulary knowledge at five years of age – over and above other language-related tests such as sound mimicry. Repetition performance also directly predicts children's ability to learn new words. Children with high scores in repetition tests are more able to learn new names than children with low scores in repetition tests (Gathercole & Baddeley, 1990). Finally, nonword repetition is used as a key marker of language impairment (e.g., Bishop, North & Donlan, 1996; Conti-Ramsden, Botting & Faragher, 2001; Ellis Weismer, Tomblin, Zhang, Buckwalter, Chynoweth & Jones, 2000), with language impaired children often having smaller vocabularies than age matched counterparts (e.g. Dollaghan & Campbell, 1998; Munson, Kurtz & Windsor, 2005).

Nonword repetition research

Nonword 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 nonword 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 nonword length increased could be explained by the decay of items in working memory.

However, subsequent research has shown that children's existing lexical and sub-lexical knowledge plays a major role in their nonword repetition ability. Gathercole (1995) re-analysed the nonwords in their original test (Gathercole & Baddeley, 1989) by separating them into "wordlike" and "non-wordlike" nonwords based on adult ratings of wordlikeness. She found that children performed significantly better for nonwords that were wordlike. Although wordlikeness is a subjective measure, even when more objective measures are used, there are still clear differences between nonwords that share substantial lexical and sub-lexical features with words compared to those that do not. For example, if one actively distinguishes nonwords based on their constituent phoneme combinations – having one set that contain highly frequent combinations of sounds versus one set containing relatively infrequent combinations – there are clear performance differences, with children regularly finding high frequency nonwords easier to repeat (e.g. Edwards, Beckman & Munson, 2004; Vitevich, Luce, Charles-Luce & Kemmerer, 1997).

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

Explanations of nonword repetition performance

Since nonword 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 and colleagues (e.g. Gathercole, 2006; Gathercole & Adams, 1993; Gathercole, Frankish, Pickering & Peaker, 1999) explained this interaction using the idea of phonological frames. Phonological working memory is used to store linguistic stimuli (e.g. nonwords in the repetition test) and when these items decay, long-term lexical and sub-lexical knowledge is relied upon to help bolster the decaying representations in working memory. Since nonwords 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 nonwords) 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 spoken (phonemic) 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 phonological 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 phonological 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 nonwords 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 nonwords containing low frequency sequences. A similar explanation can be used for wordlike versus non-wordlike nonwords. The former, since they bear great resemblance to words, are likely to be represented within the model using fewer chunks than non-wordlike nonwords.

The current paper

EPAM-VOC has thus far been used to successfully simulate the nonword results of Gathercole and Baddeley (1989), plus a nonword set devised for children between 2 and 5 years of age. Using the interaction between phonological working memory and long-term lexical and sub-lexical knowledge outlined above, the model showed a good match to children's nonword repetition performance. However, the interaction itself was not tested in detail because neither of the nonword tests were specifically designed to vary in their lexicality. Since research has shown that nonword repetition is strongly influenced by the similarity between nonwords and lexical and sub-lexical knowledge, testing children on nonwords that vary in their lexicality will increase our knowledge of the interplay between phonological working memory and long-term lexical and sub-lexical knowledge. By comparing the children's data to the results of EPAM-VOC, this research will also provide a stringent test of the interaction mechanisms employed within the model. The current paper therefore examines both child and model repetition performance across three sets of nonwords that vary in the degree of their lexicality.

The predictions from both the phonological frame account and EPAM-VOC should be that as nonwords increase in their lexicality, repetition accuracy increases. However, since repetition performance in the model is subject to the interplay between naturalistic input, phonological working memory and long-term lexical and sub-lexical knowledge, accurate predictions will arise when the simulations are performed.

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 nonword 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, the results are discussed in light of the findings presented.

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).

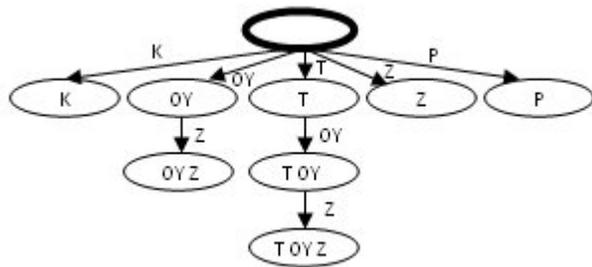


Figure 1: Graphical representation of EPAM-VOC having been presented with “Toys” (T OY Z) twice as input. Chunks are represented by ellipses and links are represented by arrows. Note that although only five phonemes are represented as single phoneme chunks (K, OY, T, Z and P) the model knows all phonemes 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 (e.g. Cowan, 1997). Within EPAM-VOC, activation is based on time – since there is 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 also 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 $(7 \times 400 \text{ ms}) + (1 \times 460 \text{ ms}) = 2,560 \text{ ms}$. 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. 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 of articulation than chunks that are high frequency. EPAM-VOC maintains a frequency for each chunk based on the number of times that the chunk has been accessed. Correct articulation of an input sequence (e.g. a nonword) 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 the additional input outlined above (reflecting the input 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 nonword 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 nonword in a nonword repetition test: 12 mothers * 10 simulations runs * 10 attempts at each nonword = 1,200 repetitions of each nonword.

5-6 year-old children’s repetition performance

Participants

18 5-6 year-old children (5;4-6;6, M=6;0; 7 male, 11 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). The nonwords in this test contained either singleton consonants only between each vowel (e.g. **rubid**) or contained at least one consonant cluster between a vowel (e.g. **glistow**). There were 5 nonwords of each type at each of three lengths (2-4 syllables). 5-syllable nonwords were excluded because children at this age had difficulty in repeating nonwords of this length. This test was considered high in lexicality since many of the nonwords included syllables that were either real words (e.g. **thickery**) or morphemes (e.g. **glistening**).

The nonword test of Dollaghan, Biber and Campbell (1995) that contained 3-syllable nonwords. 6 nonwords contained a syllable that was a lexical item (e.g. **bathesis**) and 6 changed one phoneme in the nonword so that it was entirely non-lexical (e.g. fathesis). Children show greater repetition accuracy for nonwords containing a lexical item than nonwords that do not contain a lexical item (Dollaghan, Biber & Campbell, 1995). This test was considered medium in lexicality since half of the nonwords contained lexical items and half did not.

A new nonword repetition test that was entirely non-lexical and non-morphological that contained two sets of 3-syllable nonwords: 8 that were low frequency (e.g. latmunoz) and 8 that were very low frequency (e.g. wegnerterk). The nonwords 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 nonwords 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 high lexicality nonword set had two independent variables: nonword type (single or clustered) and nonword length (2, 3, or 4 syllables). The medium lexicality nonwords had one independent variable (lexical-item: present or absent). The new nonword test also had one independent variable (frequency: low or very low). The dependent variable in all cases was the percentage of nonwords accurately recalled.

Procedure

All children were tested individually on a one-to-one basis in a quiet area of their school. Each nonword repetition test was carried out on a separate day. For the high lexicality nonwords, the original recordings were maintained (to be comparable to other studies using these nonwords); for the medium and low lexicality tests, the nonwords were recorded by a speaker native to Nottingham. The instructions for each set of nonwords are given below, and were the same for each nonword 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 simulations (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 the 120 runs for each mother are all based on a similar set of input data. The selection of one single simulation run from each mother was pseudo-random. Each set of 120 simulation runs from each mother was narrowed down to those that approximated the average of all 1,200 runs when taken as a whole. One run from each mother was randomly selected from this reduced set.

Figure 2 shows the children's results for the high lexicality nonwords together with the results from EPAM-VOC. Correlation analyses show that the model has a very similar pattern of repetition performance to the children, for both the 1,200 simulations ($r(4) = .94, p = .005$) and the 12 simulations ($r(4) = .96, p = .003$). Root Mean Square Error (RMSE) was also computed between the model and child data. RMSE reveals the average percentage difference between the model and the children. RMSE rates indicated that on average, the repetition accuracy of the 1,200 simulations was within 4.23% of the children, and the repetition accuracy of the 12 simulations was within 5.74% of the children. Taken

together, the correlations and RMSE figures show that the model provides an extremely good match to the children's data for the high lexicality nonwords.

A 2 (nonword-type: single or clustered) x 3 (nonword-length: 2, 3, or 4 syllables) repeated measures ANOVA was performed on the children's data. There was a significant effect of nonword-type ($F(1,17) = 39.62, p < .001, \eta_p^2 = .70$), showing that performance was better for single consonant nonwords, and a significant effect of nonword-length ($F(2,34) = 17.26, p < .001, \eta_p^2 = .50$), showing that performance was better for short nonwords. There was no interaction between the two ($F(2,34) = 1.53, p = .23, \eta_p^2 = .08$). For the model, there was also the same effect of nonword-type ($F(1,11) = 5.50, p = .04, \eta_p^2 = .33$) and nonword length ($F(2,22) = 9.80, p = .001, \eta_p^2 = .47$), with no interaction between the two ($F(2,22) = .33, p = .72, \eta_p^2 = .03$).

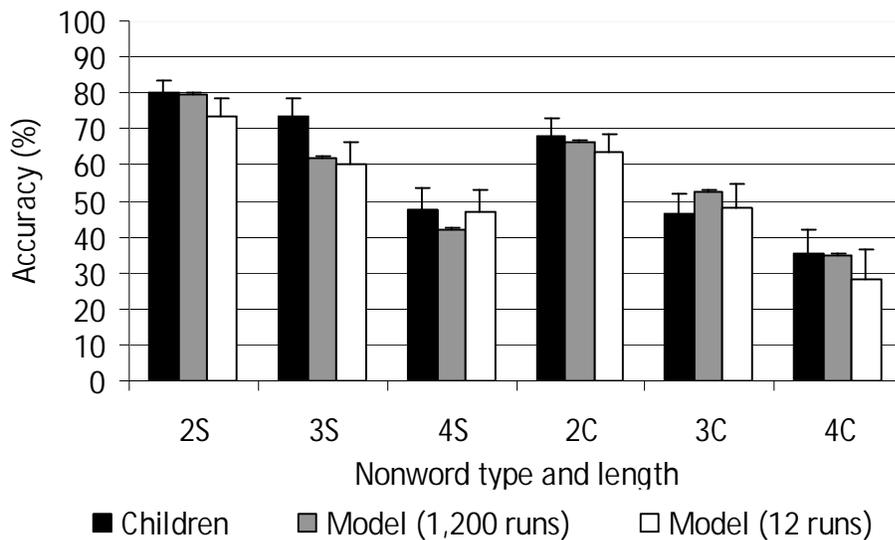


Figure 2: Nonword repetition performance (%) for the high lexicality nonwords for both the children and the two sets of model runs. The number on the x-axis denotes the number of syllables (2, 3, or 4) and the letter denotes the nonword type (S=single consonant, C=clustered consonant). Error bars indicate standard error.

Figure 3 shows the children's and model's performance for the medium and low lexicality nonwords. The children's data correlate extremely well with both the 1,200 simulations ($r(2) = .99, p = .013$) and the 12 simulations ($r(2) = .92, p = .079$). RMSE rates indicate that on average, the repetition accuracy of the 1,200 simulations was within 3.95% of the children, and the repetition accuracy of the 12 simulations was within 3.67% of the children. The correlations and RMSE figures show that the model also provides an extremely good match to the children's data for the medium and low lexicality nonwords.

For the medium lexicality nonwords, the children showed no difference in their ability to repeat nonwords containing and not containing a lexical item ($t(17) = .00, p > .99$). The same was found in the model ($t(11) = .48, p = .64$). For the low lexicality nonwords, there was no difference in children's repetition accuracy between low and very low frequency nonwords ($t(24) = 1.44, p = .17$). Again, the same result was found in the model ($t(11) = .56, p = .59$).

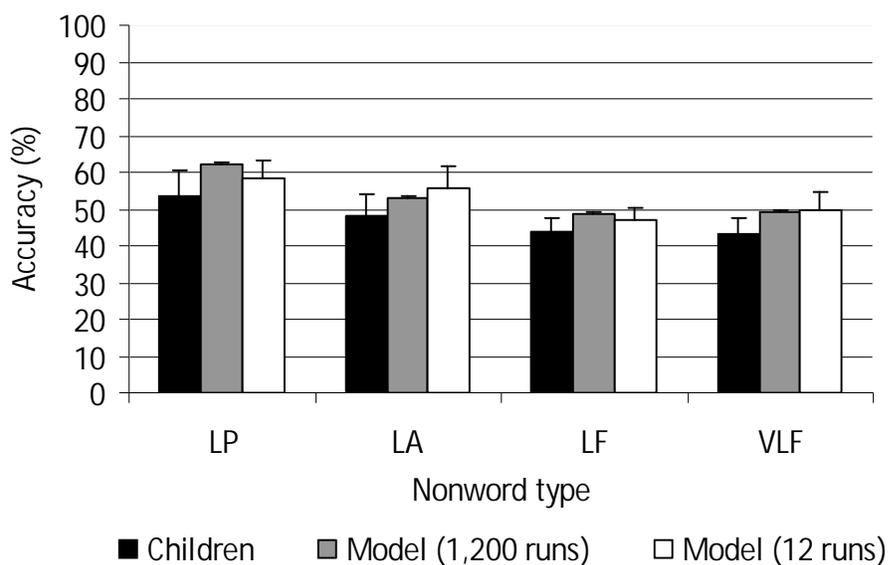


Figure 3: Nonword repetition performance (%) for the medium (leftmost two blocks) and low lexicality (rightmost two blocks) nonwords, for the children and the two sets of model runs. LP=Medium

lexicality, lexical item present; LA=Medium lexicality, lexical item absent; LF=Low lexicality, low frequency; and VLF=Low lexicality, very low frequency.

In order to examine whether there were any performance differences across nonwords that varied in their lexicality, the 3 syllable high lexicality nonwords were compared to the medium and low lexicality nonwords (both of these sets consisted of 3 syllable nonwords only). For the children, a repeated-measures one-way ANOVA (lexicality: high, medium, or low) showed no main effect of lexicality, although the results approached significance ($F(2,34) = 3.23, p = .052, \eta_p^2 = .16$). For the model, there was also no effect of lexicality ($F(2,22) = 1.38, p = .27, \eta_p^2 = .11$).

Discussion

Children's nonword repetition performance was examined across three nonword tests varying in their lexicality. For high lexicality nonwords, results were consistent with those found in 4 and 5 year old children (e.g. Gathercole & Baddeley, 1989): performance improves for single consonant nonwords and for nonwords 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). For medium lexicality nonwords, there was no difference in performance between nonwords that contained a lexical item and nonwords that did not, a finding that is inconsistent with Dollaghan, Biber and Campbell (1995) who showed that 10 year old children are more able to repeat nonwords containing a lexical item. For low lexicality nonwords, there was no difference in performance between low and very low frequency nonwords. Similar findings have been shown in the recall literature, where children's recall rates show no difference for low and very low frequency items (Gathercole et al., 1999). It is perhaps the case, therefore, that frequency effects do not distinguish between low and very low frequencies.

A model of vocabulary acquisition, EPAM-VOC, was presented that explicitly defined the interaction between phonological working memory and long-term lexical and sub-lexical knowledge

during the learning of phoneme sequences. Performance of the model showed a remarkably close match to the children for all three nonword tests. Not only were there high correlations and low RMSE rates across the data, the model also showed the same statistical effects as the children. EPAM-VOC shows how a theoretically supported account of the interaction between phonological working memory and long-term knowledge is able to capture the effects seen across three different nonword repetition tests *within the same version of the model*. The differential effects seen across nonword sets are solely because of the interplay between phonological working memory and long-term phonological chunks, without the need for any variation in parameter settings.

Let us now consider two surprising findings within both the child and model data: the lack of any effect of lexicality and the lack of any difference in performance for the two types of medium lexicality nonwords (i.e. nonwords that contain a lexical item or do not). The phonological frame account predicted that repetition accuracy would increase as the lexicality of the nonwords increased. Similarly, the account would also predict better performance for nonwords containing a lexical item over those that do not. EPAM-VOC was able to simulate both of the surprise findings seen in the children. Why is this the case? An examination of the model's performance for the medium lexicality nonwords shows that the lexical items within nonwords that contain a lexical item (e.g. 'bath' in bathesis) do not have robust representations within the model. For example, 'bath' only exists as a chunk in 78% of the model runs, and those chunks only have an average frequency of 103. The lexical items are therefore not robust enough in terms of their frequency of use to cause improved performance for nonwords containing a lexical item. The simple reason for the lack of effects across the medium lexicality nonwords, therefore, is that children of 5 and 6 years of age do not have robust representations of the lexical items within the nonwords. A similar account could explain why the children and the model do not show lexicality effects: the lexical items within the high lexicality nonwords also do not have robust representations. EPAM-VOC is therefore not only able to simulate the performance of the children, but is also able to

put forward explanations of the children's behaviour. It would be interesting to present the model with further input (e.g. based on the type of input received by older children) to see at what point lexical effects emerge.

In summary, EPAM-VOC replicates the findings of 5-6 year-old children on three different nonword repetition tests varying in the degree of their lexicality, using one solitary version of the model. 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

The author would like to thank Sarah Watson for her help in the data collection phase of this study, Marco Tamburelli for his help in the construction of the new set of nonwords outlined, Gabrielle Le Geyt for the spoken versions of the nonwords, Fernand Gobet and Julian Pine for helpful discussions on the model, and the schools and children who participated in the nonword repetition study. This research was funded by The Leverhulme Trust under grant reference F/01 374/G.

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.
- Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, *37*, 391-403.

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.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for Specific Language Impairment (SLI). *Journal of Child Psychology and Psychiatry*, 42, 741-748.

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 nonword repetition. *Applied Psycholinguistics*, 16, 211-222.

Dollaghan, C. A., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, 41, 1136-1146.

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

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J.G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Hearing, and Language Research*, 43, 865-878.

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). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513-543.

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

Gathercole, S. E., & Adams, A-M. (1993). Phonological working memory in very young children. *Developmental Psychology*, 29, 770-778.

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.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, 81, 439-454.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 84-95.

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

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. (2007). Linking working memory and long-term memory: A computational model of the learning of novel sound patterns. *Developmental Science*, 10, 853-873.

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*, 32, 1148-1176.

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.