RUNNING HEAD: LEARNING NOVEL SOUND PATTERNS

Manuscript under review at Developmental Science

Linking working memory and long-term memory: A computational model of the learning of novel sound patterns

Gary Jones[1], Fernand Gobet[2], and Julian M. Pine[3]

[1]Psychology Department, Nottingham Trent University, Burton Street, Nottingham, NG1 4BU, UK

[2]School of Social Sciences and Law, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK

[3]School of Psychology, Bedford Street South, Liverpool, L69 7ZA

Send correspondence about this article to:

Gary Jones

Psychology Department

Nottingham Trent University

Burton Street

Nottingham NG1 4BU

UK

E-mail: gary.jones@ntu.ac.uk

Telephone: +44 (115) 848-5560

Abstract

The nonword repetition (NWR) test has been shown to be a good predictor of children's vocabulary size. NWR performance has been explained using the working memory model and specifically the phonological loop, which is seen as being critical in the learning of sound patterns. However, no detailed link between long-term memory and incoming sound patterns has been proposed. A computational model of children's vocabulary acquisition (EPAM-VOC) is presented that concretely specifies how working memory and long-term memory interact. In this model performance differences arise from differences in long-term knowledge. The model's behaviour is compared with that of children in a new study of NWR, conducted in order to ensure the same nonword stimuli and methodology across ages. It is found that EPAM-VOC showed a pattern of results similar to that of children: performance is better for shorter nonwords and for wordlike nonwords, and performance improves with age. EPAM-VOC also simulates the superior performance for single consonant nonwords over clustered consonant nonwords found in previous NWR studies. EPAM-VOC represents a good approximation of the learning of novel sound patterns that specifies how working memory and long-term memory interact, using an account that indicates that capacity differences are not necessary to simulate developmental change.

Keywords: EPAM, working memory, long-term memory, nonword repetition, vocabulary acquisition, developmental change.

Introduction

Children's vocabulary learning begins slowly but rapidly increases – at the age of sixteen months children know around 40 words (Bates et al., 1994) yet by school age children learn up to 3,000 words each year (Nagy & Herman, 1987). There are individual differences across children in terms of how quickly they acquire vocabulary, and in terms of how many words they know. One of the sources of these individual differences is hypothesised to be the phonological loop component of working memory (e.g., Gathercole & Baddeley, 1989), which is perceived to be a bottleneck to the learning of sound patterns. According to this view, children with a high phonological working memory capacity are able to maintain more sound patterns and are therefore able to learn words more quickly than their low phonological working memory capacity counterparts.

The nonword repetition (NWR) test has been shown to be a reliable indicator of phonological working memory capacity and of vocabulary size. The NWR test (Gathercole, Willis, Baddeley & Emslie, 1994) involves saying a nonword to a child and asking them to speak aloud the nonword they heard. By using nonsense words, the test guarantees that the child has never heard the particular sequence of phonemes before, so there is no stored phonological representation of the nonword in the mental lexicon (Gathercole, Hitch, Service & Martin, 1997). Repeating nonwords should therefore place more emphasis on phonological working memory than on long-term phonological knowledge, and provide a more sensitive measure of phonological working memory than traditional tests such as digit span.

There are now a plethora of studies that indicate that NWR performance is the best predictor of children's vocabulary size over and above traditional memory tests such as digit span, and tests of linguistic ability such as reading tests (e.g., Gathercole

& Adams, 1993, 1994; Gathercole & Baddeley, 1989, 1990; Gathercole, Willis, Emslie & Baddeley, 1992). Furthermore, the role of phonological working memory in NWR performance is shown in adults with a specific deficit in phonological working memory, who have difficulty in learning word-nonword pairs but show no impairment for word-word pairs (e.g., Baddeley, Papagno & Vallar, 1988).

The strong relationship between NWR performance and vocabulary size led Gathercole and colleagues to hypothesise that phonological working memory plays a pivotal role in novel word learning (e.g., Gathercole & Adams, 1993; Gathercole & Baddeley, 1989; Gathercole, Willis, Baddeley & Emslie, 1994). More specifically, the phonological loop was believed to mediate the long-term storage of phonological knowledge (Gathercole & Baddeley, 1989). This was supported by Gathercole, Willis, Emslie and Baddeley (1991), who compared the influence of the phonological loop, in terms of nonword length, versus the influence of vocabulary knowledge, in terms of grammatical morphemes in a nonword. Whereas increases in nonword length led to a decline in NWR performance, the number of grammatical morphemes in a nonword had no reliable effect on NWR performance, suggesting a significant role for phonological working memory in novel word learning.

However, it is not just phonological working memory that influences NWR performance. Gathercole (1995) found that repetition performance for nonwords that were rated as wordlike was significantly better than performance for nonwords rated as non-wordlike. The implication is that long-term memory of phonological structures also influences NWR performance, and hence that there must be some form of interaction between long-term memory (LTM) and phonological working memory for NWR performance. This is supported by the fact that NWR performance significantly correlates with performance in learning word-nonword pairs, but not word-word pairs, whereas vocabulary knowledge significantly correlates with both types of pairing (Gathercole, Hitch, Service & Martin, 1997). This suggests NWR performance may only be a predictor for novel words, while vocabulary knowledge influences all types of word learning.

While the importance of LTM in the production of nonwords has been noted, it is not known exactly how phonological memory and LTM combine as yet (Gathercole, Willis, Baddeley & Emslie, 1994). Gathercole and colleagues hypothesise there to be a reciprocal relationship between phonological working memory and existing vocabulary knowledge (e.g., Gathercole, Hitch, Service & Martin, 1997), and together with the learning of novel sound patterns, the three share a highly interactive relationship (Baddeley, Gathercole & Papagno, 1998). Novel sound patterns are represented in phonological working memory but can be supported by phonological "frames" that are constructed from existing phonological representations in long-term memory (Gathercole & Adams, 1993; Gathercole, Willis, Emslie & Baddeley, 1991). Frames may contain parts of stored lexical items that share phonological sequences with the novel sound pattern contained in phonological working memory. Wordlike nonwords share more similarity with existing lexical items, resulting in better performance for wordlike nonwords over non-wordlike nonwords. Similarly, the more "novel" a sound pattern is, the more reliance will be placed on phonological working memory for learning that sound pattern.

An alternative though similar view is that it is lexical structure that influences nonword repetition performance. Metsala (1999) suggests that a child's vocabulary growth influences lexical restructuring, with words that have a large neighbourhood requiring more restructuring than those that have a sparse neighbourhood. Neighbourhood is defined as how many other words can be formed by the substitution, addition or deletion of one phoneme in the word. Words with large neighbourhoods should have an advantage over words with sparse neighbourhoods when performing phonological awareness tasks, because large neighbourhood words have been structured at a deeper level. Metsala (1999) showed that this is indeed the case. Moreover, further regression analyses showed that phonological awareness scores contributed unique variance in vocabulary size after nonword repetition scores had been entered into the regression, which was not the case when nonword repetition scores were added after phonological awareness scores. That is, lexical structure (as measured by phonological awareness tasks) was a better predictor of vocabulary size than NWR performance.

Similar less-specified theoretical positions than Gathercole and Metsala exist. For example, Munson and colleagues (e.g., Munson, Edwards & Beckman, 2005; Munson, Kurtz & Windsor, 2005) suggest that phonological representations are increasingly elaborated with age, and this would explain why performance differences in wordlike versus non-wordlike nonwords are more pronounced in younger children. Bowey (1996) argues for a phonological processing ability whereby phonological representations develop as vocabulary increases. According to this view, differences between children with high scores on NWR tests and children with low scores on NWR tests may reflect differences in their phonological processing ability rather than differences associated with phonological working memory.

Although all of these explanations indicate contributions of existing phonological knowledge and/or phonological working memory capacity, none specify how sound patterns are learned and how they interact with phonological memory. Furthermore, they do not detail how novel sound patterns are stored in LTM or how they are stored in phonological working memory. Such definitions are crucial for understanding how novel sound patterns are learned. Similarly, a precise definition of how the representations in LTM interact with those in phonological working memory is crucial in order to understand the relative roles that LTM and phonological working memory play in the learning of novel sound patterns.

The goal of this paper is to fill in this theoretical gap by providing a detailed specification of the mechanisms that link phonological working memory and LTM. We present a computational model that is able to simulate the NWR data. Not only is the model consistent with the explanations of the link between long-term and phonological working memory that have been proposed by Gathercole and Metsala, but it also fills in the detail which their explanations lack. In particular, we show that while phonological working memory is a bottleneck to language learning, LTM is more likely to be the driving force behind the learning of novel sound patterns.

The layout of the remainder of the paper is as follows. First, a summary of the existing NWR findings is given, together with a summary of existing models of NWR performance. Second, an explanation of the computational model is given. Third, a new experiment on NWR performance is presented, because existing studies do not use the same nonwords across ages, meaning that a developmental account of the model cannot be compared to the same datasets. Fourth, it is shown that the model can account for children's data in our experiment, and that the same model provides a good account of the existing NWR data. Finally, a general discussion of the findings and the model is presented.

The nonword repetition test: Existing data and simulations

There are four empirical phenomena that any computational model of NWR performance needs to simulate. First, repetition accuracy is poorer for long nonwords than it is for short nonwords. For example, Gathercole and Baddeley (1989) found that 4-5 year old children's NWR performance was superior for 2-syllable nonwords than 3-syllable nonwords, and for 3-syllable nonwords than 4-syllable nonwords respectively. Second, children's repetition accuracy gets better with age. For example, Gathercole and Adams (1994) found 5 year olds' NWR performance to be superior to that of 4 year olds. Third, performance is better for single consonant nonwords than clustered consonant nonwords (e.g., Gathercole & Baddeley, 1989). Fourth, NWR performance is better for wordlike nonwords than it is for non-wordlike nonwords, suggesting the involvement of LTM representations of phoneme sequences (Gathercole, 1995).

Two influential models of nonword repetition exist, although neither was created with the intention of accounting for the key phenomena listed above. Hartley and Houghton (1996) detail a connectionist network that is presented with nonword stimuli in the training phase and is tested on the same nonwords in a recall phase. Decay incorporated within the model means that longer nonwords are recalled with less accuracy than shorter nonwords. Furthermore, the model is able to simulate certain types of error in nonword repetition. For example, the phonemes in a syllable have competition from other related phonemes such that substitutions can take place. Based on data from Treiman and Danis (1988), the model displays similar types of error to those made by children and adults.

Brown and Hulme (1995, 1996) detail a trace decay model in which the incoming list of items (e.g., nonwords) is represented as a sequence of 0.1 second time slices. For example, a nonword may take 0.5 seconds to articulate and will therefore comprise 5 segments, or 5 time slices of 0.1 seconds each. Each segment can vary in strength from 0 to 1, with segments beginning with a strength of 0.95

when they enter memory. As time progresses (i.e., every 0.1 seconds), each segment of the input is subjected to decay. For example, an item that occupies 5 segments will enter memory one segment at a time, and thus the first segment of the item will have been subjected to four periods of decay by the time the fifth segment of the item enters memory. Decay also occurs when the item is being articulated for output. To combat items decaying quickly, the strength of certain items is increased based on relationships to LTM traces, such that, for example, wordlike nonwords would increase in strength more than non-wordlike nonwords.

Long nonwords decay more quickly than short nonwords, thereby supporting the existing literature on children's performance of nonword repetition where repetition accuracy gradually decreases across 2 to 4 syllables. This leads to the prediction that long words will take longer for children to acquire than short words, and this prediction seems to be borne out by age-of-acquisition data (Brown & Hulme, 1996).

In terms of the four criteria outlined at the beginning of this section, both models can account for longer nonwords being repeated back less accurately than shorter nonwords. However, none of the other criteria was simulated within either model. Furthermore, neither model explains how sound patterns are actually learned through exposure to naturalistic stimuli. A computational model will now be presented that (a) details how novel sound patterns are learned, (b) explains how these sound patterns then interact with phonological working memory, and (c) accounts for the key phenomena we have described.

A new computational model of nonword repetition: EPAM-VOC

EPAM (Feigenbaum & Simon, 1984) is a computational architecture that progressively builds a discrimination network of knowledge based on the input it receives. The discrimination network is hierarchical such that at the top there is a root node, below which several further nodes will be linked. Each of these nodes may in turn have further nodes linked below them, creating a large and organised knowledge base of the input received. Visually, the resulting hierarchy of nodes and links can be seen as a tree, and indeed EPAM shares similarities to what are known in computer science as “trie” structures (Fredkin, 1960).

EPAM and its variants have been used to model human performance in various psychological domains, such as learning, memory, and perception in chess (De Groot & Gobet, 1996; Gobet, 1993; Gobet & Simon, 2000; Simon & Gilmartin, 1973), verbal learning behaviour (Feigenbaum & Simon, 1984), the digit-span task (Richman, Staszewski & Simon, 1995), the context effect in letter perception (Richman & Simon, 1989), and the acquisition of syntactic categories (Freudenthal, Pine & Gobet, 2002, 2006; Gobet, Freudenthal & Pine, 2004; Jones, Gobet & Pine, 2000a) (see Gobet et al., 2001, for an overview). Thus, most of the mechanisms used in the model described in this paper have been validated by independent empirical and theoretical justifications, and their validity has been established in a number of diverse domains.

The hierarchical structure of EPAM is particularly suited to the learning of sound patterns. If one considers a sentence, it can be broken down into a sequence of phonemes that represent each of the words in the sentence. This sequence of phonemes needs to be stored in a hierarchical fashion to preserve the order of the phonemes. EPAM provides a simple mechanism by which this can be accomplished, such that the resulting discrimination network becomes a long-term memory of sound patterns. Preliminary versions of the models have been described in Jones, Gobet and Pine (2000b, 2005). This section will first describe how EPAM-VOC builds a discrimination network of sound patterns, and second, how phonological memory will be simulated and linked to the discrimination network.

Learning sound patterns in EPAM-VOC

The standard EPAM architecture builds a hierarchy of nodes and links that exist as a cascading tree like structure. EPAM-VOC is a simplified version of EPAM that uses phonemic input in order to build a hierarchy of phonemes and sequences of phonemes.

When a sequence of phonemes is presented, EPAM-VOC traverses as far as possible down the hierarchy of nodes and links. This is done by starting at the top node (the root node) and selecting the link that matches the first phoneme in the input. The node at the end of the link now becomes the top node and EPAM-VOC tries to match the next phoneme from the input to all the links below this node. If an appropriate link exists, then the node at the end of the link becomes the top node and the process repeats. When a node is reached where no further traversing can be done (e.g., the next phoneme does not exist in the links below the current top node, or the node has no links below it), then learning occurs in one of two ways:

1. Phoneme learning. If the next phoneme in the input sequence does not exist as a link below the root node, the phoneme is added as a link and node below the root node such that EPAM-VOC now has knowledge of the phoneme.
2. Sequence learning. If the next phoneme in the input sequence exists as a link below the root node, it is added as a link and node below the current top node. As a result, a sequence of phonemes is learned consisting of the phonemes that were used to traverse the network up to the current top node, plus the new phoneme

just added. Sequence learning, where increasingly larger "chunks" of phonemes are acquired, is very similar to discrimination in traditional EPAM networks.

EPAM-VOC begins with a null root node, meaning that the model begins with no knowledge of phonemes or phoneme sequences. When EPAM-VOC receives an input (a sequence of phonemes), new nodes and links are created. The initial learning for EPAM-VOC involves phoneme learning, so that each phoneme of the English language exists as a node below the root node. One may expect the child to already have knowledge of such phonemes, so it should be noted that the primary aim of phoneme learning is to ensure that all phoneme sequences beginning with a particular phoneme occur as nodes below the particular phoneme in the network. The vast majority of learning in the network will be sequence learning, where the information at nodes becomes sequences of phonemes and therefore segments of speech (e.g., specific words) rather than just individual sounds (i.e., phonemes).

Let us consider an example of the network learning the utterance "What?". Using the CMU Lexicon database, this utterance is converted to the phonemic representation "W AH1 T". Note that the phonemic input to the model does not specify gaps between words, but does specify the stress on particular phonemes (0=unstressed; 1=primary stress; 2=secondary stress).

When EPAM-VOC first sees the phonemic representation "W AH1 T", it tries to match as much of the input as possible using its existing knowledge, and then learn something about the remainder of the input. In attempting to match the input to EPAM-VOC's existing knowledge, the first part of the input ("W") is applied to all of the root node's links in the network. Since EPAM-VOC begins with no knowledge, the first time "W AH1 T" is input, EPAM-VOC tries to match "W" and fails. At this point phoneme learning takes place, because the phoneme “W” does not exist as a link below the root node. A new node is created together with a link from the root node to the new node with the test “W”. Hence the phoneme “W” is learnt and EPAM-VOC can move on to the next part of the input (“AH1 T”). Again, “AH1” cannot be matched and so EPAM-VOC learns this phoneme in a similar manner before moving on to the remainder of the input (“T”). This is also learnt in a similar manner. Thus after first encountering “W AH1 T”, EPAM-VOC learns each of the constituent phonemes in the word.

When “W AH1 T” is input to EPAM-VOC a second time, a match can be made with the first part of the input (“W”), and the “W” link can be taken such that the new top node becomes the “W” node. EPAM-VOC now moves on to the remainder of the input (“AH1 T”) and tries to match the first part of the remaining input (“AH1”) by examining the links below the current top node. Since the “W” node does not have any links below it, no further matching can take place. At this point, EPAM-VOC examines the remainder of the input and realises that it already knows about the “AH1” phoneme, and so sequence learning can occur. A new node and link is created below the “W” node containing the phoneme “AH1”. Some learning has taken place at the current top node, and so the current top node reverts back to being the root node, and EPAM-VOC moves on to the remainder of the input (“T”). This part of the input can be matched below the root node such that the “T” node becomes the current top node, but as there is no further input, no further learning takes place.

Using “W AH1 T” as input a third time, EPAM-VOC is able to match the first part of the input (“W”), and so makes the “W” node the new top node. The next part of the input is then examined (“AH1”), and because this exists as a link below the “W” node, it can be matched, with the “W AH1” node, becoming the current top node. The matching process then moves on to the next part of the input (“T”), but as no links exist below the “W AH1” node, no matching can take place. At this point, sequence learning can take place (because the phoneme “T” is already known to the model), and so a new node and link “T” can be made below the current top node. Thus after three successive inputs of the sequence “W AH1 T”, the whole word is learnt, and the network is as shown in Figure 1.

Insert figure 1 about here

This simple example serves to illustrate how EPAM-VOC works; in the actual learning phase each input line is only used once, encouraging a diverse network of nodes to be built. Note that EPAM-VOC needs to know individual phonemes before they can be learnt as part of a sequence of phonemes. For example, should the network in Figure 1 see the utterance “Which?” (“W IH1 CH”), it will traverse down the “W” link, and move on to the next part of the input (i.e., “IH1 CH”). However, the network does not know the phoneme “IH1”, and so phoneme learning needs to take place, learning the individual phoneme “IH1” before moving on to the remainder of the input “CH” (and learning this as an individual phoneme also). Although learning may seem to occur rather quickly within EPAM-VOC, it is possible to slow it down (e.g., by manipulating the probability of learning a new node), and this has been successful for other variants of EPAM models (e.g., Croker, Pine & Gobet, 2003; Freudenthal, Pine & Gobet, 2002). Reducing the learning rate is likely to yield the same results, but over a longer period of time. For the input sets that will be used here, which contain a very small subset of the input a child would hear, it is therefore sensible to have learning take place in the way that has been illustrated.

Implementing phonological memory and linking it to the discrimination network

EPAM-VOC now requires a specification of phonological memory, or more specifically, the phonological loop, and a method by which the loop interacts with EPAM-VOC's discrimination network. As in the standard working memory model (Baddeley & Hitch, 1974), the storage part of the phonological loop, the phonological store, is a decay based store which allows items to remain in the store for 2,000 ms (Baddeley, Thomson & Buchanan, 1975). EPAM-VOC therefore has a time-limited store that allows 2,000 ms of input.

In the standard working memory model, the phonological loop also has a sub-vocal rehearsal mechanism, which allows items to be rehearsed in the store such that they can remain there for more than 2,000 ms. However, Gathercole and Adams (1994) suggest that children of five and under do not rehearse, or at least if they do, they are inconsistent in their use of rehearsal. Furthermore, Gathercole, Adams and Hitch (1994) found no correlation between articulation rates and digit span scores for four year old children, suggesting that children of four years of age do not rehearse (if they did, there should be a relationship between articulation rate and digit span because rehearsal rate would be related to how quickly the child could speak words aloud). Previous computational models have also shown that it is not necessary to simulate rehearsal in order to model memory span (e.g., Brown & Hulme, 1995). Hence EPAM-VOC does not use the sub-vocal rehearsal mechanism. The input is cut off as soon as the time limit is reached (i.e., the input representations are not refreshed), and so the phonological loop becomes a phonological store, in-line with current findings regarding rehearsal in young children.

Having described the model’s LTM (i.e., the discrimination network of nodes and links) and phonological store, we are now in a position to discuss the mechanisms enabling these two components to interact. This is the central contribution of this paper, as there is currently no clear explanation in the literature as to how the phonological store links to LTM and how this relation is modulated by learning. Within EPAM-VOC, it is relatively easy to specify how sound patterns in LTM interact with the phonological store. When sound patterns are input to EPAM-VOC, they are matched to those that are stored as nodes in the discrimination network; for any sound patterns that can be matched in LTM, a pointer to the relevant node is placed in the phonological store. That is, input sounds are not necessarily stored individually in the phonological store, but are mediated by LTM nodes that contain neural instructions as to how to produce them. The amount of information that can be held in the phonological store is thus mediated by the amount of information already stored in LTM. Retrieving each node and processing each phoneme within a node requires a certain amount of time, and the cumulative time required by these processes provides an explanation of how much information can be held in the phonological store. Let us explain in detail how this works.

The length of time taken to represent the input is calculated based on the number of nodes that are required to represent the input. The time allocations are based on values from Zhang and Simon (1985), who estimate 400 ms to match each node, and 84 ms to match each syllable in a node except the first (which takes 0 ms). (These estimates are derived from adult data.) As the input will be in terms of phonemes, with approximately 2.8 phonemes per syllable (based on estimates from the nonwords in the NWR test), the time to match each phoneme in a node is 30 ms.

Consider as an example the input “What about that?” (“W AH1 T AH0 B AW1 T DH AE1 T”). Given the network depicted in Figure 1, all that can be represented in the phonological store within the 2,000 ms timescale is “W AH1 T AH0 B AW1”. The “W AH1 T” part of the input is represented by a single node, and is allocated a time of 460 ms (400 ms to match the node, and 30 ms to match each constituent item in the node excluding the first item). Most of the other phonemes are not known to the model and are assumed to take the same time as a full node (400 ms; the time allocated to each phoneme is assumed to be constant). This means that only three additional phonemes can be represented within the phonological store, by which time the actual input to the model has required a time allocation of 1,660 ms. Matching another node would cost at least 400 ms, and thus exceed the time capacity of the store. When the EPAM-VOC network is small, and nodes do not contain much information, only a small amount of the input can be represented in the phonological store. When the EPAM-VOC network is large, the model can use nodes that contain large amounts of information, and therefore a lot of the input information can be represented in the phonological store. Larger networks also enable more rapid learning, as increasingly large chunks of phonemes can be put together to create new chunks (i.e., new nodes in the discrimination network).

It is worth noting that EPAM-VOC can readily simulate phenomena from the adult literature on working memory tasks, although it was not developed with this specific aim in mind. For example, the word length effect (e.g., Baddeley, Thomson & Buchanan, 1975) can be simulated under the assumption that a word will be represented as a single node in the model. Longer words will contain more phonemes within that node and will therefore take longer to be matched. The word frequency effect (e.g., Whaley, 1978) can be simulated under the assumption that timing estimates are reduced for nodes that are accessed frequently because, with exposure, the information held in a sequence of nodes gets chunked in a single node (see Freudenthal, Pine & Gobet, 2005, for a description of how this mechanism has been used for simulating data on syntax acquisition).

How EPAM-VOC fits in with existing accounts of the link between LTM and phonological working memory

While much more detailed and specified as a computer program, the EPAM-VOC explanation of the influence of existing phonological knowledge on NWR performance is actually consistent with that suggested by Gathercole and colleagues. EPAM-VOC learns individual sounds (i.e., phonemes) and also sequences of phonemes, or mini-sound patterns, that are not themselves words. Phoneme sequences can be used to aid the remembering of unfamiliar word forms, and in particular wordlike nonwords that are more likely to match phonological sequences in LTM. The reliance on the phonological store as a mediator of verbal learning therefore depends on EPAM-VOC's existing knowledge of sound patterns, which is determined by the amount and variability of linguistic input the model receives.

EPAM-VOC is also consistent with Metsala's (1999) hypothesis surrounding neighbourhood size. EPAM-VOC learns more detail for words with large neighbourhoods relative to words with small neighbourhoods. Large neighbourhood words by definition have many other words that differ only by a single phoneme, whereas small neighbourhood words do not. All other things being equal, this means that EPAM-VOC learns more about large neighbourhood words because similar phoneme sequences will be seen as input. For example, compare the large neighbourhood word make (which has neighbours such as take and rake) with the small neighbourhood word ugly. EPAM-VOC will learn something about make even if it does not ever see the word, because if the model is shown take or rake as input, the ending phoneme sequence of these words are shared by make. On the other hand, few similar words exist for ugly and so relevant phoneme sequences are only likely to be learned by EPAM-VOC if ugly itself is presented to the model.

Existing explanations of the link between phonological knowledge and the phonological store suggest that the phonological store mediates NWR performance – it is the bottleneck to language learning (e.g., Gathercole, in press). Given that it is already known that existing phonological knowledge influences NWR performance, an alternative source of individual variation is the amount of phonological knowledge the child currently has – some children may have either been exposed to more linguistic input, more variation in linguistic input, or both. This is what will be explored in the simulations presented here. It will be shown that although phonological working memory is a bottleneck that restricts how much information can be learned, the amount of information that can fit into phonological working memory is likely to be strongly determined by children's existing phonological knowledge. It will also be shown that it is possible to explain differences in children's NWR performance across ages purely in terms of differences in the amount of knowledge of sound patterns that has built up in LTM. The implication is that developmental changes in working memory capacity are not necessary in order to explain developmental changes in children's NWR performance.

A study of nonword repetition performance

EPAM-VOC offers the opportunity to examine developmental change in NWR performance. Comparisons of NWR performance can be made between young children and the model at an early stage of its learning, and between older children and the model at a later stage of its learning. Unfortunately, the current set of NWR studies has tended to use different sets of stimuli (Gathercole, 1995), making comparison difficult. Furthermore, existing studies have carried out nonword repetition tests in different ways. For example, in Gathercole and Baddeley (1989), the children heard a cassette recording of the nonwords, whereas in Gathercole and Adams (1993), the children heard the experimenter speaking aloud the nonwords with a hand covering the speaker's mouth. This reduces the consistency of the current NWR results. We therefore decided to collect additional empirical data in order to assess children's NWR performance across ages using the same nonword stimuli and the same experimental method.

The children who participated in this experiment were of 2-5 years of age, the ages at which NWR performance correlates best with vocabulary knowledge. A pilot experiment using 1-4 syllable lengths showed that younger children had great difficulty repeating back the 4-syllable nonwords, and so nonwords of length 1-3 syllables were used across all age groups (Gathercole & Adams, 1993, used 1-3 syllable nonwords for their 2-3 year old children).

Method

Participants

There were 127 English speaking children, of which 66 were 2-3 years of age (mean = 2.49; SD = 0.47) and 61 were 4-5 years of age (mean = 4.22; SD = 0.33). All children were recruited from nurseries (2-3 year olds) and infant schools (4-5 year olds) within the Derbyshire area. Six of the 2-3 year olds and one of the 4-5 year olds failed to complete the experiment leaving 120 children in total.

Design

A 2x2x3 mixed design was used with a between-subject independent variable of age (2-3, 4-5) and within-subject independent variables of nonword type (wordlike, non-wordlike), and nonword length (1, 2, 3 syllables). The dependent variables were nonword repetition response, vocabulary score, and span score.

Materials

A set of 45 nonwords of 1, 2, and 3 syllables were constructed. Five wordlike and 5 non-wordlike nonwords were used at each syllable length based on subjective mean ratings of wordlikeness as rated by undergraduate students. The remaining nonwords were not used. Examples of wordlike and non-wordlike nonwords at each of 1, 2, and 3 syllables respectively are: dar, yit, ketted, tafled, commerant, and tagretic (the stress for all nonwords was strong for the first syllable). The full list of nonwords used can be seen in the appendix. One audiotape was created, consisting of read-aloud versions of the wordlike and non-wordlike nonwords in a randomised order (as per the methodology of Gathercole & Baddeley, 1989). The randomised order was the same for all children.

Nine different coloured blocks of equal size were used for a verbal memory span task, with three pre-determined sequences from length 2 to length 9 being created. For example, one of the sequences for length 3 was a red block, followed by a blue block, followed by a green block. A blocks task was used instead of the traditional digit span task because it was assumed that young children would be more familiar with colours than numbers.

The British Picture Vocabulary Scale (BPVS, Dunn, Dunn, Whetton & Burley, 1997) was used to establish vocabulary size.

<u>Procedure</u>

All children were tested in the first term of school. Before commencing the experiment, the researcher spent an afternoon in each school and nursery in order to familiarise themselves with the children. All children were tested individually in a quiet area of the school/nursery. The order of testing was consistent across all children: BPVS followed by NWR followed by digit span. The BPVS used difficulty level 1 for the 2-3 year olds and difficulty level 2 for the 4-5 year olds. In all cases, there were up to fourteen trials of 12 items each, with testing ending when 8 errors were made within a trial. The NWR test was carried out using an audiocassette player to present the nonwords in a randomised order. Each child was informed they would hear some "funny sounding made up words" and that they should try and repeat back immediately exactly what they heard. The experimenter noted whether the repetition was correct, partially correct (i.e., at least one phoneme correct), completely wrong, or if no response was given. For the block test, each child was given each of two sequences of coloured blocks (starting at length two). If each was repeated back correctly, then the length was increased by one and the process began again. If only one was repeated correctly, then a sequence from the third list was taken and if this was repeated back correctly, the length was increased by one and the process began again using sequences from the original two lists. Span length was taken as the highest length at which the child successfully repeated two sequences.

Results

Descriptive statistics are shown in Table 1. A 2 (age: 2-3 year old or 4-5 year old) x 2 (nonword type: wordlike or non-wordlike) x 3 (nonword length: 1, 2, or 3 syllables) ANOVA was carried out on the data. There was a significant main effect of age ($F(1,118)=201.73$, $Mse=338.94$, $p<.001$), with older children performing better on the nonword repetition test. There was also a significant main effect of nonword type ($F(1,118)=603.47$, $Mse=196.36$, $p<.001$), wordlike nonwords being repeated back more easily than non-wordlike nonwords. There was also a significant main effect of nonword length ($F(2,236)=260.52$, $Mse=116.93$, $p<.001$). Post-hoc Bonferroni tests showed that one-syllable nonwords were repeated back more easily than both two-syllable nonwords and three-syllable nonwords, and two-syllable nonwords were repeated back more easily than three-syllable nonwords (all $p<.001$). There was no interaction between age and nonword type ($F(1,118)=3.84$, $Mse=1.25$, $p>.05$), but significant interactions existed for age and nonword length ($F(2,236)=67.09$, $Mse=30.11$, $p<.001$) and nonword type and nonword length ($F(2,236)=7.53$, $Mse=2.52$, $p<.001$). There was no three-way interaction ($F(2,236)=.01$, $Mse=.01$, $p>.05$).

Insert table 1 about here

In terms of span and BPVS scores, both measures showed superior performance for the older children ($F(1,118)=113.63$, $Mse=4.50$, $p<.001$, and $F(1,118)=382.11$, $Mse=13.75$, $p<.001$, respectively). Note that these two analyses are based on log transformed scores in order to ensure homogeneity of variance.

For the 2-3 year old children, there were significant correlations between span scores and vocabulary size ($r(58)=.56$, $p<.001$) and between nonword repetition scores and vocabulary size ($r(58)=.49$, $p<.001$). While the correlation between nonword repetition and vocabulary size may seem low at first glance, this is in fact a higher correlation than the significant correlation of .34 found by Gathercole and Adams (1993).

For the 4-5 year old children, there were significant correlations between span scores and vocabulary size ($r(58)=.81$, $p<.001$) and between nonword repetition scores and vocabulary size ($r(58)=.76$, $p<.001$).

Discussion

The present results are consistent with existing NWR studies: children's performance declines as the length of the nonword increases; children's NWR performance is better for wordlike rather than non-wordlike nonwords; and older children perform better at repetition than their younger counterparts. The results also clarify an anomaly in previous NWR literature, where children's nonword repetition performance was better for two-syllable nonwords than it was for one-syllable nonwords. Here, the reverse is true – children perform better on one-syllable nonwords than on all other lengths of nonword (as was found by Roy & Chiat, 2004). This supports the explanation put forward by Gathercole and Baddeley themselves that there were problems with the one-syllable nonwords they used (Gathercole & Baddeley, 1989).

The correlational data are also consistent with previous findings, where significant correlations have been found between nonword repetition performance and vocabulary size, and between span scores and vocabulary size. Children with high NWR scores tend to have a larger vocabulary, as do children with high span scores.

The basic nonword repetition results and the results of the correlational analysis show a high degree of consistency with previous studies of nonword repetition, establishing a solid base for guiding the computer simulations.

Simulating the nonword repetition results

Carrying out the NWR test

The NWR test for the model consisted of presenting each nonword as input to the model and seeing if it could represent the nonword within the 2,000 ms time capacity. However, children's NWR performance is clearly error prone, whereas EPAM-VOC currently has no opportunity to make errors, except for being unable to fit a nonword into the store within the time limitation. Using one-syllable nonwords as an example, and assuming all necessary phonemes are known to the model, EPAM-VOC would fit all one syllable nonwords into the store because they have a maximum of three phonemes – even if each phoneme was only matched as a single node in the network, the allocated time capacity would still only be 1,200 ms (3*400 ms).

An error-producing mechanism was therefore introduced whereby EPAM-VOC could probabilistically take an incorrect link while traversing the network. Thus, EPAM-VOC now produces repetition errors even when all phonemes can fit into the phonological store. The probability of producing an error was decreased as more input was seen by the model (see Table 2), because it was assumed that as children get older, they become more adept at encoding and articulating the sound patterns they receive.

After the model has seen 25% of the input, the probability of taking an incorrect link was set at .10. This figure was not arbitrary but reflected the error rates in 2-3 year old children. In our experiment, single-syllable error rates were 24% and 50% for wordlike and non-wordlike nonwords, respectively; in Gathercole and Adams's (1993) study, the corresponding error rates were 17% and 22% for words and nonwords, respectively. This averages at an error rate of 28%. The average length of all the one-syllable words and nonwords used by the two studies is 3.1 phonemes. A word or nonword of 3 phonemes would normally require three traversals to nodes in the network (one for each phoneme). If each traversal has a probability of error of .10, then the probability of making a correct sequence of three traversals is .9*.9*.9=.73, or a 27% error rate, which closely matches the 28% average error rate for single-syllable words and nonwords. Although the error rate was set to match that of one syllable items, the same was not true for two and three syllable items where the rate of error was open to the dynamics of the model.

At the end of the input, the probability of making a traversal error was assumed to decrease to .04. Thus, for single-syllable nonwords, the probability of making a correct sequence of traversals was .96*.96*.96=.88, which corresponds to a 12% error rate.

The input regime

The simulations used both mothers' utterances and pairs of random dictionary words as input. The utterances were taken from the Manchester corpus (Theakston, Lieven, Pine & Rowland, 2001), which includes twelve sets of mother-child interactions between mothers and 2-3 year olds recorded over a one year period. The average number of utterances for each mother was 25,519 (range 17,474-33,452). Pairs of random dictionary words were selected from the CMU Lexicon database (available at http://www.speech.cs.cmu.edu/cgi-bin/cmudict). Pairs of words were used in order to keep consistent the number of phonemes used as input – the average number of phonemes in an utterance (across all mother's) was 12.03; the average number of phonemes in a random word from the CMU Lexicon database was 6.36.

The relative ratio of mother's utterances and pairs of random words from the lexicon were gradually altered to reflect an increased variation in input as the child grows older. The first 25% of the mother's input was seen by EPAM-VOC, and thereafter gradually more and more pairs of random lexicon words were included within that input.

Table 2 shows, at each stage of the model's learning, the exact values that were used for the proportion of mother's utterances to pairs of lexicon words. In terms of input, EPAM-VOC was presented with the same number of utterances that appeared in the mother's corpus, but some of these were replaced by pairs of random lexicon words based on the amount of pairs of lexicon words that should be included in the input. For example, Anne's mother used 31,393 utterances in total. At the beginning, EPAM-VOC was presented with the first 25% of these utterances, but for the next 12.5% of the utterances, every tenth utterance was replaced with a pair of random lexicon words (to reflect the 10% of pairs of random lexicon words that needed to be input to the model, as indicated in Table 2). At this point, if a nonword repetition test was carried out, there would be a .09 probability of traversing down an incorrect link.

Insert table 2 about here

Although comparisons to the child data will only be made at certain points in the model's learning (to correspond to 2-3 and 4-5 year old children), EPAM-VOC will be examined later at each developmental stage of learning in order to illustrate exactly how it was able to simulate the child data.

For all simulations, all input was converted into a sequence of phonemes using the CMU Lexicon database. This database cross-references words with the phonemic form of each word. All of the phonemes used in the database map onto the standard phoneme set for American English. The use of phonemic input assumed that some form of phonemic feature primitives already existed to distinguish one phoneme from another, which would be expected for children of two years and above. The phonemic input did not distinguish word boundaries, so no word segmentation had been performed on the input that is being fed to the model.

Simulations of the data

A total of 120 simulations were carried out (ten for each of the sets of mother's utterances). Ten simulations per set of utterances were used in order to produce a robust set of results, given that the model has a random component (the possibility of selecting an incorrect link when traversing the network for matching nonwords). Changes to the input and the probability of making a traversal error were incorporated in accordance with the values in Table 2. Nonword repetition results were averaged across the 120 simulations.

To compare EPAM-VOC with 2-3 year old children's NWR performance, an NWR test was taken after the model had seen 25% of the input (i.e., when only mother's utterances had been seen as input). To compare EPAM-VOC with 4-5 year old children, an NWR test was taken after EPAM-VOC had seen 87.5% of the input.

Descriptive statistics are shown in Table 1. A 2 (stage of learning: early [25% of input] or late [87.5% of input]) x 2 (nonword type: wordlike or non-wordlike) x 3 (nonword length: 1, 2, or 3 syllables) ANOVA was carried out on the data. There was a significant main effect of stage of learning ($F(1,238)=495.60$, $Mse=490.0$, $p<.001$), with the late model performing better on the nonword repetition test. There was also a significant main effect of nonword type ($\underline{F}(1,238)=63.30$, $\underline{Mse}=76.54$, $\underline{p}<.001$), wordlike nonwords being repeated back more easily than non-wordlike nonwords. There was also a significant main effect of nonword length ($\underline{F}(2,476)=310.98$, $\underline{Mse}=314.86$, $\underline{p}<.001$). Post-hoc Bonferroni tests showed that one-syllable nonwords were repeated back more easily than both two-syllable and three-syllable nonwords, and two-syllable nonwords were repeated back more easily than three-syllable nonwords (all $\underline{p}<.001$). There was an interaction between stage of learning and nonword type ($\underline{F}(1,238)=18.20$, $\underline{Mse}=22.00$, $\underline{p}<.001$), between nonword type and nonword length ($\underline{F}(2,476)=35.32$, $\underline{Mse}=34.95$, $\underline{p}<.001$), and between stage of learning and nonword length ($\underline{F}(2,476)=42.48$, $\underline{Mse}=43.01$, $\underline{p}<.001$). There was also a significant three-way interaction ($\underline{F}(2,476)=6.40$, $\underline{Mse}=6.34$, $\underline{p}<.01$).

Figure 2 shows a comparison between early EPAM-VOC and the 2-3 year old children and Figure 3 shows a comparison between late EPAM-VOC and the 4-5 year old children. When all data-points for the model were correlated with those of the children, there was a highly significant correlation ($\underline{r}(10) = .91$, $\underline{p} < .001$; RMSE=9.08).

Insert figure 2 about here

Insert figure 3 about here

The pattern of effects in NWR performance for EPAM-VOC is very similar to that of the children in the experiment presented earlier in this paper. Of even more importance is the fact that the results fit in with three of the four key criteria outlined earlier. First, nonword repetition performance declines as nonword length increases. Second, repetition performance improves at later stages in the model’s learning. Third, wordlike nonwords have a better repetition accuracy than non-wordlike nonwords.

However, although the new data provided a solid base on which to test the model, the experiment did not include single and clustered consonant nonwords, which was the fourth criterion that must be met by any computational model of nonword repetition. In order to show that EPAM-VOC also fulfils this criterion, the model will be compared to the single and clustered consonant NWR performance of the four and five year olds used by Gathercole and Baddeley (1989). Two additional NWR tests were carried out using the nonwords used by Gathercole and Baddeley (their nonwords can be seen in the appendix). To compare with four year olds, a NWR test was taken after the model had seen 75% of the input, and to compare to five year olds, a NWR test was taken after the model had seen 100% of the input. The amounts of input fit in with the 87.5% level that was used when comparing 4-5 year olds in the study presented in this paper. Note that because of the problems outlined earlier regarding the one-syllable nonwords used in the Gathercole and Baddeley (1989) study, these are omitted from the analysis.

Figure 4 shows the repetition performance for single consonant nonwords for EPAM-VOC at 75% and 100% of the model’s learning, and 4 and 5 year old children and Figure 5 shows the repetition performance for clustered consonant nonwords for EPAM-VOC at 75% and 100% of the model’s learning, and for 4 and 5 year old children. When all data-points for the model were correlated with those of the children, there was a highly significant correlation ($r(10) = .89$, $p < .001$; RMSE=14.94).

A 2 (stage of learning: 75% of input or 100% of input) x 2 (nonword type: single or clustered) x 3 (nonword length: 2, 3 or 4 syllables) ANOVA was carried out on the data. There was a significant main effect of stage of learning ($F(1,238)=75.61$, $Mse=69.34$, $p<.001$), with repetition performance being better for the 100% model. There was also a significant main effect of nonword type ($F(1,238)=27.78$, $Mse=30.04$, $p<.001$), with better repetition performance for single consonant nonwords over clustered consonant nonwords. There was also a significant main effect of nonword length ($F(3,714)=898.79$, $Mse=849.16$, $p<.001$). Post-hoc Bonferroni tests showed that two-syllable nonwords were repeated back more easily than both three-syllable and four-syllable nonwords, and three-syllable nonwords were repeated back more easily than four-syllable nonwords (all $p<.001$). There were no two-way or three-way interactions (all $p>.05$). The pattern of effects on repetition performance is consistent with that found by Gathercole and Baddeley (1989). In particular, an important result of this section is that EPAM-VOC, like children, performs better for single consonant nonwords than for clustered consonant nonwords.

Insert figure 4 about here

Insert figure 5 about here

Although repetition errors have not been analysed in great detail in children's NWR studies, it has been noted that, for example, the highest proportion of errors in five year olds is due to phonological substitution (Gathercole, Willis, Baddeley & Emslie, 1994). In the study presented, the nonwords were not recorded and therefore we have no data regarding the types of error that the children made. However, an analysis of the types of error made by the model showed that 64% of errors were phonological substitutions, 22% were phonological additions, and 11% were phonological deletions. Phoneme additions/deletions/substitutions were defined as two or less phonemes being added/deleted/substituted within a nonword. The model's tendency to make substitution errors is a direct consequence of the model's mechanism for simulating production errors, which involves (occasionally) taking incorrect links when traversing the network.

Summary of the simulations

EPAM-VOC provided a very good match to the new data from the experiment presented here, and the model also showed the same pattern of results that were seen in the 4 and 5 year old children studied by Gathercole and Baddeley (1989), although the goodness of fit was perhaps not as pronounced in this case as that obtained with the new data. The main issue for the 4 and 5 year old comparisons was that the model had a rather low repetition accuracy for four-syllable nonwords. This suggests that perhaps EPAM-VOC had not seen enough input or enough variation in input. The problem for the model, given that variation in the input is critical, is in determining the type and amount of input that a 4 or 5-year-old child is likely to have heard. Clearly, this is a very difficult task and any attempt to replicate the input is likely to be a crude approximation. For example, even though the model received half of the mother's utterances as input, this only constituted 3,046 different words on average. The lexicon words were used as an attempt to bolster this amount, but they are just an approximation of the diversity of input that 4 and 5 year old children receive. The model thus provided a good approximation of existing repetition performance based on what would seem to be a reasonable, but not perfect, approximation of the input. The results suggest that using more realistic input is likely to result in an even better match to the data.

How EPAM-VOC simulates nonword repetition

Thus far, it has been shown that EPAM-VOC, in spite of its relative simplicity, accounts for the NWR findings surprisingly well. How does EPAM-VOC achieve such a good fit to the results? Let us again turn to the four criteria outlined in the introduction, which specified what a model of NWR performance must be able to achieve. These will be considered in turn, and an explanation given for how EPAM-VOC satisfies each of them.

NWR performance is better for short nonwords than long nonwords

In EPAM-VOC, longer nonwords are less likely to be represented in full within the phonological store until the model has learnt a lot about sound patterns, and so the model has difficulty repeating longer nonwords during the early stages of its learning. This can be illustrated by examining the time that is required to represent nonwords at various stages of the model's learning. Figure 6 shows the average time to represent non-wordlike nonwords at different stages of the model's input (averaged across all 120 simulations). The figure clearly shows that for short nonwords, there is little benefit to further learning, as the model masters repetition of these nonwords at an early stage. For longer nonwords, however, mastery occurs at a much later stage as EPAM-VOC learns more about the phonemic input and is therefore able to represent the nonwords using fewer nodes than at earlier stages.

Insert figure 6 about here

NWR performance improves with age

A further illustration of how the model improves with more learning can be shown by plotting the number of nodes that are learnt at various stages of learning. Figure 7 shows that such a plot is almost linear. However, it should be pointed out that learning at later stages involves nodes that contain large sequences of phonemes, rather than nodes that contain short sequences of phonemes, which are what is found early on in learning. Performance thus improves with age because more knowledge about sequences of phonemes is learnt as EPAM-VOC receives more input – and this means that EPAM-VOC is more able to fit longer nonwords within the time limit of the phonological store.

Insert figure 7 about here

NWR performance is better for single consonant than clustered consonant nonwords

Improved performance for single consonant nonwords over clustered consonant nonwords is actually very easy to explain once one considers the number of phonemes required to articulate each type of nonword. The single consonant nonwords used by Gathercole and Baddeley (1989) contain an average of 5.50 phonemes whereas the clustered consonant nonwords contain an average of 7.75 phonemes. Children are therefore likely to find clustered consonant nonwords more difficult to repeat back because these nonwords are, in effect, longer. Similarly, in EPAM-VOC, it will be more difficult to fit clustered consonant nonwords into the phonological store than single consonant nonwords.

NWR performance is better for wordlike than non-wordlike nonwords

There is a slight difference in the phonemic length of wordlike nonwords and non-wordlike nonwords because non-wordlike nonwords tend to have clustered consonants. There is an average of 5.00 phonemes for wordlike nonwords versus 5.67 phonemes for non-wordlike nonwords in the experiment presented. This in itself is unlikely to be sufficient to produce such striking performance differences between the two types of nonword. In terms of the model, wordlike nonwords are expected to contain phoneme sequences that are more familiar (i.e., that exist in already known words) than non-wordlike nonwords. Assuming that these sequences occur frequently in the input, EPAM-VOC should learn a substantial number of them, and therefore the component phonemes in wordlike nonwords should be stored as larger sequences of phonemes than the component phonemes in non-wordlike nonwords. Hence, what is expected is that wordlike nonwords can be represented using fewer nodes than non-wordlike nonwords, meaning they can be represented in less time within the phonological store. Subjecting the model's performance to the same ANOVA reported previously, but using the time to match nonwords as the dependent measure rather than nonword repetition scores, shows a highly significant difference for the type of nonword ($F(1,216)=844.26$, $Mse=7.74$, $p<.001$ [log transformed data]), with non-wordlike nonwords taking longer to be represented within the phonological store. It is clear that wordlike nonwords can be represented using fewer nodes than non-wordlike nonwords, and this is why it takes these nonwords less time to be matched in the phonological store.

General discussion

In the last decades, short-term memory capacity has been measured in two ways. Starting with Miller (1956), one group of researchers have proposed that STM capacity can be measured in chunks, that is, perceptual units. This idea has been embodied in EPAM, an influential computational model of perception, learning, and memory that has been applied to a number of domains ranging from chess expertise to letter recognition. Another group of researchers, centred around Baddeley and Hitch's (1974) model of working memory, have proposed that the capacity of short-term memory – in particular auditory short-term memory – is time-based. Building on work by Zhang and Simon (1985) with adults, this paper has shown that these two approaches can be reconciled. In particular, we have shown that important data on phonemic learning can be explained by a computational model, EPAM-VOC, that (a) incrementally builds up chunks of knowledge about phonological sequences in long-term memory, and (b) specifies the relation between working memory and long-term memory.

Using the NWR task as a test-bed, we identified four criteria that any viable model should meet. The simulations presented in this paper have demonstrated that EPAM-VOC fulfils all of these criteria via an interaction between a fixed capacity phonological store and the chunking of phonemic knowledge, together with variation in the amount of input. First, repetition accuracy was poorer for long nonwords than it was for short nonwords, which fits the children's data on NWR performance (e.g., Gathercole & Adams, 1993; Gathercole, Willis, Emslie & Baddeley, 1991) and the findings of the experiment presented here. Second, repetition accuracy improved at each stage of the model's learning, mirroring the fact that, as children grow older, their NWR accuracy improves (e.g., Gathercole, 1995; Gathercole & Adams, 1994). Third, performance was better for single consonant nonwords than clustered consonant nonwords, which is consistent with the findings of Gathercole and Baddeley (1989). Fourth, NWR performance was better for wordlike nonwords than it was for non-wordlike nonwords, which is supported both in previous literature (e.g., Gathercole, 1995; Gathercole, Willis, Emslie & Baddeley, 1991) and in the new experiment of NWR performance presented here.

In addition to simulating the NWR data very well, EPAM-VOC makes two important theoretical contributions. First, it concretely specifies how phonological working memory interacts with existing LTM phonological knowledge. Second, the simulations illustrate how differences in performance across ages do not require explanations based around capacity differences – rather, the explanation is based on the extent of existing phonological knowledge. We expand on these contributions in turn.

Interaction of phonological working memory with LTM knowledge

The explanation of how phonological working memory interacts with LTM knowledge is both parsimonious and elegant. The model gradually builds up a discrimination network of phonological knowledge in order to increase the amount of information that can be held in the phonological store. As input is received by the model, any existing long-term representations of any part of the input can be accessed such that if the model knows a three phoneme sequence, for example, those three phonemes do not need to be stored individually within the phonological store but rather a pointer can be stored to the equivalent node containing the sequence. As a result, the more phonological knowledge the model has in its LTM, the more items can be stored in the phonological store. Precisely how the phonological loop interacts with LTM has never been defined before in computational terms.

While more precise and quantitative, EPAM-VOC's account fits in with current views of how phonological working memory and LTM interact. Gathercole and colleagues (e.g., Gathercole & Adams, 1993; Gathercole, Willis, Emslie & Baddeley, 1991) propose that phonological working memory is supported by phonological "frames" that are constructed from existing phonological representations in long-term memory. EPAM-VOC is able to operationalise this description: phonological frames are phonological sequences, and the way in which they interact with phonological working memory is captured by the idea that an input is recoded into sequences as much as possible. Wordlike nonwords share more phonological sequences with real words (which will have been learnt from the input) and so they have an advantage over non-wordlike nonwords that share less similarity with real words. In this way, EPAM-VOC predicts, as Gathercole and colleagues also predict, that the more "novel" a sound pattern is, the more reliance is placed on phonological working memory when learning it.

Metsala (1999) hypothesises that it is the segmental structure of items in LTM that is critical for performance in nonword repetition. Wordlike nonwords are repeated more accurately than non-wordlike nonwords because wordlike nonwords have more lexical neighbours, and so they can be represented using larger lexical units. This is exactly what is found in the EPAM-VOC simulations where the nodes (i.e., the existing phoneme sequences in the EPAM-VOC network) that are used to represent wordlike nonwords are larger than those that are used to represent non-wordlike nonwords (because wordlike nonwords are more likely to share phoneme sequences with real words). This means that wordlike nonwords can be represented using fewer nodes than non-wordlike nonwords. Furthermore, Metsala found that children of 4-5 years of age showed better performance for early acquired words than later acquired words in onset-rime blending tasks – a finding that would be predicted by EPAM-VOC under the assumption that the model will have more detailed nodes for early acquired words, because they are likely to have occurred more frequently in the input.

The key concept for Metsala (1999) is that it is vocabulary growth that influences lexical restructuring. Words having large neighbourhoods require more restructuring than words with sparse neighbourhoods, and thus there is more lexical structure surrounding large neighbourhood words. The difference between this view and that implemented in EPAM-VOC is that there is no restructuring in EPAM-VOC – learning reflects a deeper level of structure rather than restructuring per se. Nevertheless, both accounts are able to explain performance on nonword repetition tests without using phonological working memory as the primary influence.

Are capacity differences necessary for explaining performance differences?

EPAM-VOC has shown that children's NWR performance can be simulated without the need for variations in capacity. Gathercole, Hitch, Service and Martin (1997) suggested that the capacity of the phonological loop is influenced by two factors – a "pure" capacity that differs across individuals and with development/maturation, and the amount of vocabulary knowledge held at any one time. The results presented here suggest that capacity differences are not necessary, at least to explain developmental changes in NWR performance. Capacity differences have often been cited in the developmental literature yet it is actually difficult to measure capacity size without tapping into some form of long-term knowledge. For example, the digit span task is often used as a test of "pure" capacity; yet, it relies on children's long-term knowledge of digits and digit sequences – and hence the NWR test has been found to be a purer test of phonological working memory capacity (e.g., Gathercole & Adams, 1993). This paper has shown that even the NWR task may suffer from the same problem.

The difficulty of measuring memory capacity limitations is well known, especially in domains where learning is continuous (Lane, Gobet & Cheng, 2001), and other computational models have also questioned whether capacity differences produce the best explanation of the children's data. For example, Jones, Ritter and Wood (2000) found that differences in strategy choice rather than capacity provided the best explanation of children's problem solving performance.

Some developmental theorists have also denied the role of memory capacity per se. For example, Case (1985) suggests that children have a functional memory capacity. In much the same way as in EPAM-VOC, as task experience increases, more complex knowledge structures can be held in memory, leading to improved task performance. EPAM-VOC can therefore be seen as an operationalised version of the Case theory that is focused on the task of language learning. Moreover, there is no reason to suggest that the same mechanisms used by EPAM-VOC could not be applied to other developmental tasks. For example, Chi (1978) and Schneider, Gruber, Gold and Opwis (1993) examined children's chess playing, finding that working memory capacity for chess-based information increased as a function of expertise, yet for other tasks, such as digit span, no difference was found between the chess players and controls. The mechanisms presented in this paper suggest that children's chess expertise leads them to have a deeper structuring of chess knowledge in their LTM, and this facilitates how much information they can hold in WM in much the same way as EPAM-VOC's network of sound patterns facilitates the amount of input that can be processed within its phonological store.

Further predictions of the model

The process by which LTM and phonological working memory interact in EPAM-VOC makes specific predictions regarding children's and adult's language capabilities. First, children who have more phonological knowledge in LTM should perform better on NWR tasks. An obvious follow-on from this is that, children who perform better on NWR tasks should, in turn, be more productive in their language use. This is exactly what was found by Adams and Gathercole (2000), who showed that four year old children who performed well on NWR tests produced a greater number of unique words and also produced longer utterances than children who performed less well on the NWR tasks. In line with the mechanisms proposed in this paper, good performance on NWR tasks is indicative of an above average knowledge base for phonological sequences, which is suggestive of a larger vocabulary. In turn, an above average knowledge base would mean the existence of large sequences of phonemes in LTM, and therefore the child being able to produce longer utterances within the same capacity.

Second, children and adults who are multi-lingual should be able to perform better on NWR tasks because they have a comparatively larger amount of phonological knowledge in LTM. Multi-lingual speakers have learnt sound patterns for two or more languages and thus their phonological knowledge is likely to be much richer than their monolingual counterparts. There are already studies that provide support for this prediction.

Papagno and Vallar (1995) found that adult polyglots (defined by them as people who were fluent in at least three languages) performed better on NWR tasks than non-polyglots. The same findings have been found in children (Masoura & Gathercole, 2005). In fact, the findings of Masoura and Gathercole are strongly predicted by EPAM-VOC. Masoura and Gathercole split Greek children learning English into low and high vocabulary groups (based on vocabulary performance in English-Greek translation tests) and low and high NWR groups (based on NWR performance for English and Greek nonwords). EPAM-VOC would predict that any differences on English word learning tests would be governed by vocabulary knowledge, and hence differences should only be seen between the low and high vocabulary groups. This is exactly what Masoura and Gathercole found.

Conclusion

EPAM-VOC represents an important step not only in the simulation of NWR performance but also in the definition of working memory and how it links to LTM. The way in which EPAM-VOC links short-term and long-term memory is such that at an early stage of the model’s learning, emphasis is placed on short-term memory (in this case, the phonological store). At later stages of the model’s learning, emphasis is placed on long-term memory. The architecture of EPAM-VOC is consistent with the idea that task experience is critical in order to process as many items as possible within a store of limited duration and capacity. With limited or no task experience, very few items can be processed in short-term memory and thus short-term memory acts as a bottleneck to long-term learning. With more task experience, increasingly large amounts of information can be processed in short-term memory, which in turn allows more opportunity for further information to be learnt. The beauty of this architecture is that developmental differences that are often attributed to capacity changes can arise solely through exposure to a task – under the assumption that young children have less exposure to developmental tasks than their older counterparts. That is, apparent developmental changes in capacity arise from relative experience with components of the task at hand.

EPAM-VOC is obviously only a first attempt at simulating the learning of novel sound patterns. There are clearly areas where the model is limited. For example, relationships between phonemes are not represented, such that phenomena such as the phonological similarity effect (e.g., Conrad & Hull, 1964) cannot be simulated. However, improvements to the model could be made by considering further findings in the vocabulary acquisition and memory literature, and considering other computational models in this area (e.g., Burgess & Hitch, 1992).

The model presented here represents a good first pass at learning novel sound patterns. The model is able to simulate nonword repetition findings surprisingly well, and provides important insights into the sophistication of the child language learner. EPAM-VOC is the first step in modelling vocabulary acquisition using a parsimonious model and using large-scale datasets as input.

EPAM-VOC reconciles time-based and chunked-based approaches to memory capacity. By doing so, it provides well-specified mechanisms on the relation between working memory and long-term memory, in particular explaining how long-term knowledge interacts with working memory limitations. These mechanisms shed light not only on how the bottleneck imposed by limitations on working memory restricts learning ability, but also on how the capacity of this bottleneck changes as a function of what has been learned. The implication is that developmental changes in performance on working memory tasks may be an indirect effect of increases in underlying knowledge rather than a direct effect of changes in the capacity of working memory.

References

Adams, A-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. International Journal of Language and Communication Disorders, 35, 95-116.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. Psychological Review, 105, 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (pp. 47-90). New York, NY: Academic Press.

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long term learning depends on short-term storage. Journal of Memory and Language, 27, 586-595.

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

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. Journal of Child Language, 21, 85-123.

Bowey, J. A. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. Journal of Experimental Child Psychology, 63, 44-78.

Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed? Journal of Memory and Language, 34, 594-621.

Brown, G. D. A., & Hulme, C. (1996). Nonword repetition, STM, and word age-of-acquisition: A computational model. In S. E. Gathercole (Ed.), Models of short-term memory (pp. 129-148). Hove, UK: Psychology Press.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. Journal of Memory and Language, 31, 429-460.

Case, R. (1985). Intellectual development: Birth to adulthood. Orlando, Florida: Academic Press.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J: Erlbaum.

Conrad, R., & Hull, A. (1964). Information, acoustic confusion, and memory span. British Journal of Psychology, 55, 429-432.

Croker, S., Pine, J. M., & Gobet, F. (2003). Modelling children's negation errors using probabilistic learning in MOSAIC. In F. Detje, D. Dörner & H. Schaub (Eds.) Proceedings of the Fifth International Conference on Cognitive Modeling (pp. 69-74). Bamberg:Universitäts-Verlag.

De Groot, A. D., & Gobet, F. (1996). Perception and memory in chess. Heuristics of the professional eye. Assen: Van Gorcum.

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

Fredkin, E. (1960). Trie memory. Communications of the ACM, 3, 490-499.

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

Freudenthal, D., Pine, J. M. & Gobet, F. (2002). Modelling the development of Dutch optional infinitives in MOSAIC. In W. D. Gray & C. D. Schunn (Eds.), Proceedings of the 24th Annual Meeting of the Cognitive Science Society (pp. 328-333). Mahwah, NJ: Erlbaum.

Freudenthal, D., Pine, J. M., & Gobet, F. (2005). Resolving ambiguities in the extraction of syntactic categories through chunking. Cognitive Systems Research, 6, 17-25.

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

Gathercole, S. E. (in press). Nonword repetition and word learning: The nature of the relationship. Applied Psycholinguistics.

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. & Adams, A-M. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. Journal of Memory and Language, 33, 672-688.

Gathercole, S. E., Adams, A-M., & Hitch, G. J. (1994). Do young children rehearse? An individual differences analysis. Memory & Cognition, 22, 201-207.

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., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. Developmental Psychology, 33, 966-979.

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.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1991). The influence of number of syllables and wordlikeness on children's repetition of nonwords. Applied Psycholinguistics, 12, 349-367.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. Developmental Psychology, 28, 887-898.

Gobet, F. (1993). A computer model of chess memory. In W. Kintsch (Ed.), Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society (pp. 463-468). Boulder, CO: Erlbaum.

Gobet, F., Freudenthal, D., & Pine, J. M. (2004). Modelling syntactic development in a cross-linguistic context. In W. G. Sakas (Ed.), Proceedings of the COLING 2004 Workshop "Psychocomputational Models of Human Language Acquisition" (pp. 53-60). Geneva: COLING.

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.

Gobet, F. & Simon, H.A. (2000). Five seconds or sixty: Presentation time in expert memory. Cognitive Science, 24, 651-682.

Hartley, T. & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. Journal of Memory and Language, 35, 1-31.

Jones, G., Gobet, F. & Pine, J.M. (2000a). A process model of children's early verb. In L.R. Gleitman & A.K. Joshu (Eds.), Proceedings of the Twenty-Second Annual Meeting of the Cognitive Science Society (pp. 723-728). Mahwah, NJ: Erlbaum.

Jones, G., Gobet, F. & Pine, J. M. (2000b). Learning novel sound patterns. In N. Taatgen & J. Aasman (Eds.), Proceedings of the 3rd International Conference on Cognitive Modeling (pp.169-176). Groningen, Netherlands: Universal Press.

Jones, G., Gobet, F. & Pine, J.M. (2005). Modelling vocabulary acquisition: An explanation of the link between the phonological loop and long-term memory. Artificial Intelligence and Simulation of Behaviour Journal, 1, 509-522.

Jones, G., Ritter, F.E. & Wood D.J. (2000). Using a cognitive architecture to examine what develops. Psychological Science, 11, 93-100.

Lane, P. C. R., Gobet, F., & Cheng, P. C. H. (2001). What forms the chunks in a subject's performance? Lessons from the CHREST computational model of learning. Behavioral and Brain Sciences, 24, 128-129.

Masoura, E. V. & Gathercole, S. E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. Memory, 13, 422-429.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. Journal of Educational Psychology, 91, 3-19.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Munson, B., Edwards, J., & Beckman, M. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. Journal of Speech, Language, and Hearing Research, 48, 61-78.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. Journal of Speech, Language, and Hearing Research, 48, 1033-1047.

Nagy, W. E. & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), The nature of vocabulary acquisition (pp. 19-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Papagno, C. & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. Quarterly Journal of Experimental Psychology, 48A, 98-107.

Richman, H. B., & Simon, H. A. (1989). Context effects in letter perception: Comparison of two theories. Psychological Review, 3, 417-432.

Richman, H. B., Staszewski, J., & Simon, H. A. (1995). Simulation of expert memory with EPAM IV. Psychological Review, 102, 305-330.

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: Evidence from typically developing children. Journal of Speech, Language, and Hearing Research, 47, 223-234.

Schneider, W., Gruber, H., Gold, A., & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. Journal of Experimental Child Psychology, 56, 328-349.

Simon, H. A., & Gilmartin, K. J. (1973). A simulation of memory for chess positions. Cognitive Psychology, 5, 29-46.

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.

Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. Journal of Experimental Psychology: Learning, Memory and Cognition, 14, 145-152.

Whaley, C. P. (1978). Word-nonword classification time. Journal of Verbal Learning and Verbal Behavior, 17, 143-154.

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

Appendix

Nonwords used in the study presented, with phonemic representations

Wordlike nonwords

DAR	(D AA1 R)
LAN	(L AE1 N)
FOT	(F AO1 T)
TULL	(T AH1 L)
DUTT	(D AH1 T)
JARDON	(JH AA1 R D AH0 N)
DINNULT	(D IH1 N AH0 L T)
KETTED	(K EH1 T AH0 D)
RINNER	(R IH1 N ER0)
LITTING	(L IH1 T IH0 NG)
VOLERING	(V AA1 L ER0 IH1 NG)
COMMERANT	(K AA1 M ER0 AE1 N T)
BANNAFER	(B AE1 N AE1 F ER0)
HAPPAMENT	(HH AE1 P AH0 M AH0 N T)
CANNARRATE	(K AE1 N EH1 R EY2 T)

Non-wordlike nonwords

GICK	(G IH1 K)
FOLL	(F AA1 L)
JID	(JH IH1 D)
DOP	(D AA1 P)
YIT	(Y IH1 T)
MOSTER	(M AO1 S T ER0)
LIGDALE	(L IH1 G D EY1 L)
HAGMENT	(HH AE1 G M AH0 N T)
PUNMAN	(P AH1 N M AE1 N)
TAFLED	(T AE1 F L EH1 D)
DOPPELRATE	(D AA1 P AH0 L R EY1 T)
TACOVENT	(T AA1 K OW0 V EH1 N T)
DERPANEST	(D ER1 P AE1 N EH1 S T)

NARTAPISH	(N AA1 R T AH0 P IH1 SH)
TAGRETIC	(T AE1 G R EH1 T IH0 K)

Gathercole and Baddeley's (1989) nonwords, with phonemic representations

Single consonant

PENNEL	(P EH1 N AH0 L)
BALLOP	(B AE1 L AH0 P)
RUBID	(R UW1 B IH0 D)
DILLER	(D IH1 L ER0)
BANNOW	(B AE1 N OW0)
DOPPELATE	(D AO1 P EH0 L EY0 T)
BANNIFER	(B AE1 N AH0 F ER0)
BARRAZON	(B AE1 R AH0 Z AA0 N)
COMMERINE	(K AA1 M ER0 IY0 N)
THICKERY	(TH IH1 K ER0 IY0)
WOOGALAMIC	(W UW1 G AE0 L AE1 M IH0 K)
FENNERISER	(F EH1 N ER0 AY1 Z ER0)
COMMEECITATE	(K AH1 M IY1 S AH0 T EY0 T)
LODDENAPISH	(L AA1 D EH0 N EY1 P IH0 SH)
PENNERIFUL	(P EH1 N ER1 IH0 F UH0 L)

Clustered consonant

HAMPENT	(HH AE1 M P EH0 N T)
GLISTOW	(G L IH1 S T OW0)
SLADDING	(S L AE1 D IH0 NG)
TAFFLEST	(T AE1 F L EH0 S T)
PRINDLE	(P R IH1 N D AH0 L)
GLISTERING	(G L IH1 S T ER0 IH0 NG)
FRESCOVENT	(F R EH1 S K AH0 V AH0 N T)
TRUMPETINE	(T R AH1 M P AH0 T IY0 N)
BRASTERER	(B R AE1 S T ER0 ER0)
SKITICULT	(S K IH1 T AH0 K AH0 L T)
CONTRAMPONIST	(K AA1 N T R AE1 M P AH0 N AH0 S T)

PERPLISTERONK	(P ER1 P L IH1 S T ER0 AA0 NG K)
BLONTERSTAPING	(B L AA1 N T ER0 S T EY1 P IH0 NG)
STOPOGRATTIC	(S T AA1 P OW0 G R AE1 T IH0 K)
EMPLIFORVENT	(EH1 M P L IH0 F AO1 R V EH0 N T)

Table 1.

Mean correct responses (standard deviations in parentheses) for all dependent measures, for the children and EPAM-VOC. Maximum scores for the NWR, coloured block task, and BPVS tests were 5, 9, and 168, respectively.

	2-3 year olds	EPAM-VOC, 25% of input	4-5 year olds	EPAM-VOC, 87.5% of input
Wordlike nonwords, one-syllable	3.77 (.72)	3.53 (1.00)	4.47 (.57)	4.22 (.77)
Wordlike nonwords, two-syllables	3.18 (.70)	2.96 (1.10)	4.27 (.69)	3.65 (1.04)
Wordlike nonwords, three-syllables	1.78 (.87)	2.09 (1.02)	3.87 (.72)	3.47 (1.01)
Non-wordlike nonwords, one-syllable	2.53 (.68)	3.51 (1.11)	3.40 (.72)	4.22 (.87)
Non-wordlike nonwords, two-syllables	2.28 (.99)	2.38 (1.12)	3.55 (.87)	3.60 (1.05)
Non-wordlike nonwords, three-syllables	.53 (.57)	.57 (.73)	2.77 (.83)	2.88 (1.27)
Coloured block task	2.25 (.44)		3.33 (.68)	
BPVS	27.18 (5.74)		53.25 (9.42)	

Table 2.

Parameter values at each stage of the model's learning.

Amount of input seen by the model (%)	Percentage of pairs of lexicon words included in the input	Probability of selecting an incorrect link
0 – 25	0	.10
25 - 37.5	10	.09
37.5 – 50	20	.08
50 - 62.5	30	.07
62.5 – 75	40	.06
75 - 87.5	50	.05
87.5 – 100	60	.04

Figure legends

Figure 1. Structure of an EPAM-VOC net after receiving the input “W AH1 T” three times.

Figure 2. Repetition accuracy for EPAM-VOC after 25% of the input, plotted against 2-3 year old children.

Figure 3. Repetition accuracy for EPAM-VOC after 87.5% of the input, plotted against 4-5 year old children.

Figure 4. Single consonant nonword repetition accuracy for EPAM-VOC after 75% and 100% of the input, plotted against the 4 year old and 5 year old children of Gathercole and Baddeley (1989).

Figure 5. Clustered consonant nonword repetition accuracy for EPAM-VOC after 75% and 100% of the input, plotted against the 4 year old and 5 year old children of Gathercole and Baddeley (1989).

Figure 6. Average time to match non-wordlike nonwords at various stages of EPAM-VOC’s learning.

Figure 7. Nodes learned by EPAM-VOC at various stages of learning.

Figure 1

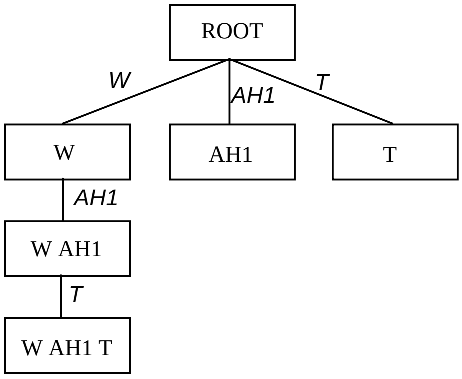


Figure 2

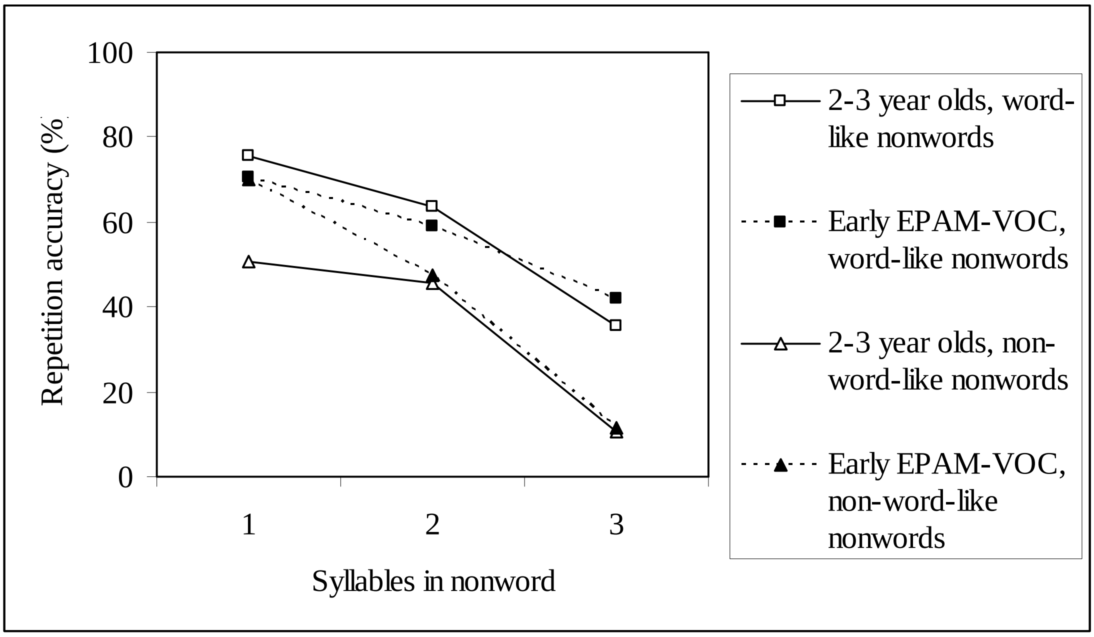

100
80
60
40
20
0
Repetition accuracy (%)
1
2
3
Syllables in nonword
2-3 year olds, word-like nonwords
Early EPAM-VOC, word-like nonwords
2-3 year olds, non-word-like nonwords
Early EPAM-VOC, non-word-like nonwords

Figure 3

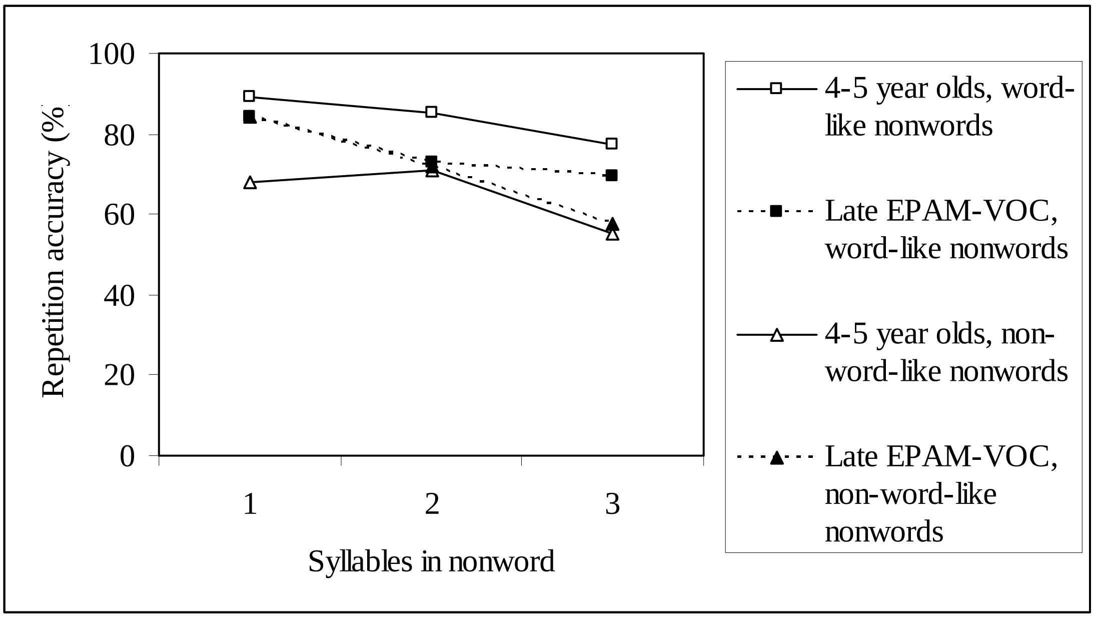

100
80
60
40
20
0
Repetition accuracy (%
1
2
3
Syllables in nonword
4-5 year olds, word-like nonwords
Late EPAM-VOC, word-like nonwords
4-5 year olds, non-word-like nonwords
Late EPAM-VOC, non-word-like nonwords

Figure 4

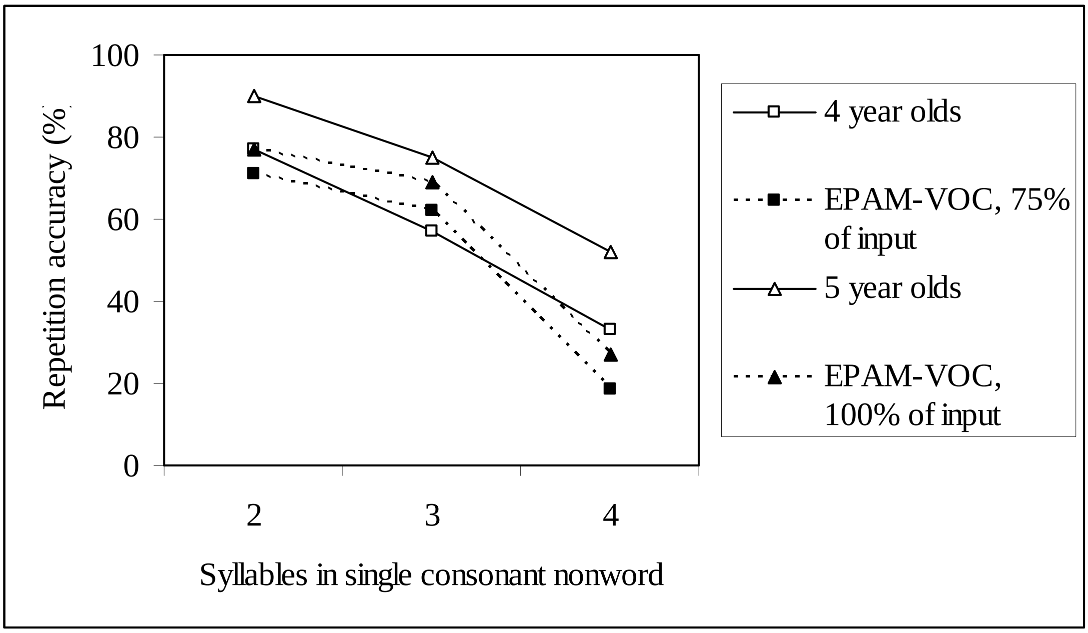

100
80
60
40
20
0
Repetition accuracy (%)
2
3
4
Syllables in single consonant nonword
4 year olds
EPAM-VOC, 75% of input
5 year olds
EPAM-VOC, 100% of input

Figure 5

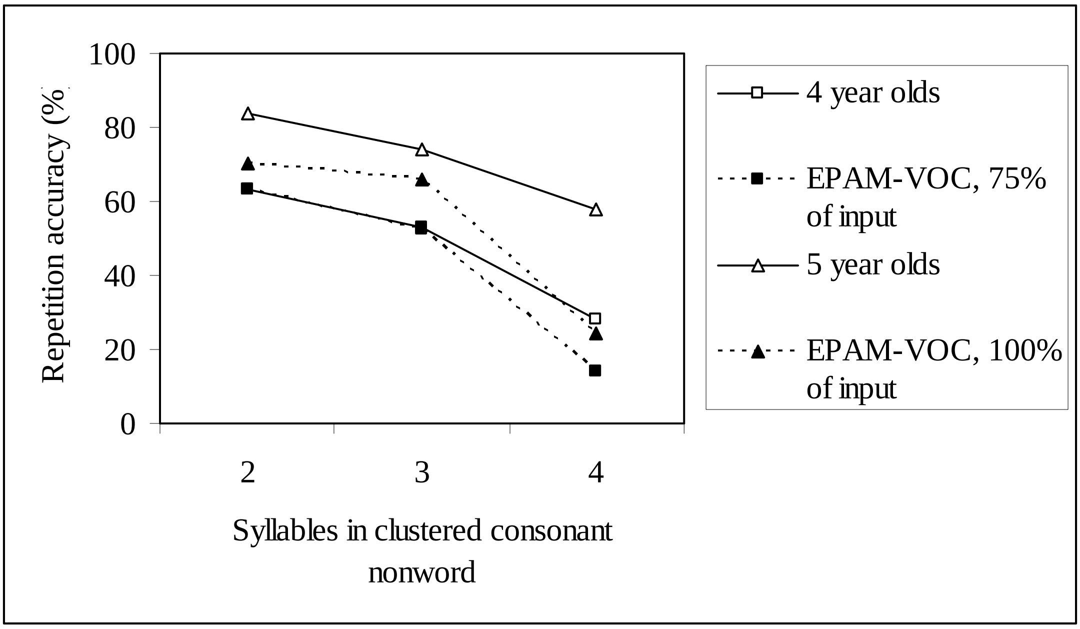

100
80
60
40
20
0
Repetition accuracy (%)
2
3
4
Syllables in clustered consonant nonword
4 year olds
EPAM-VOC, 75% of input
5 year olds
EPAM-VOC, 100% of input

Figure 6

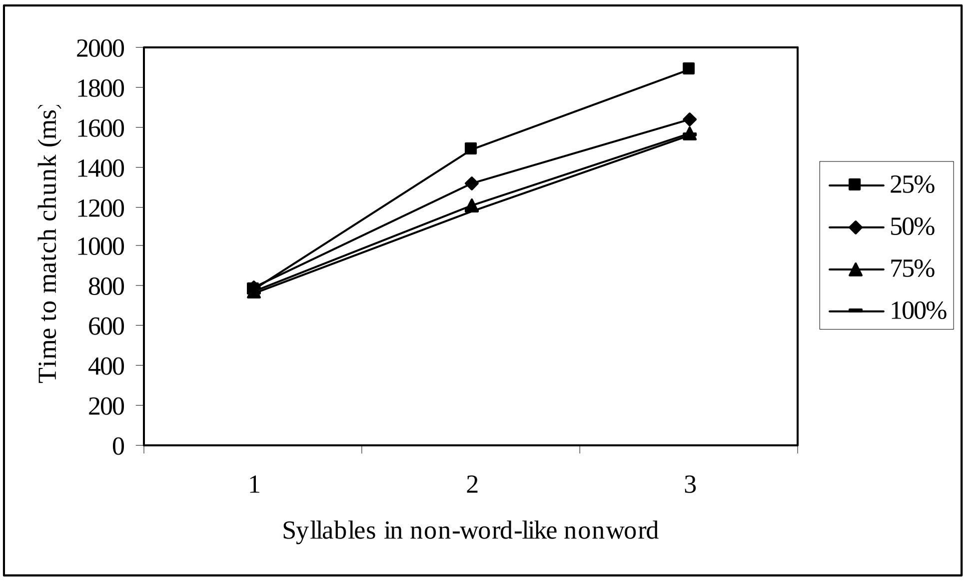
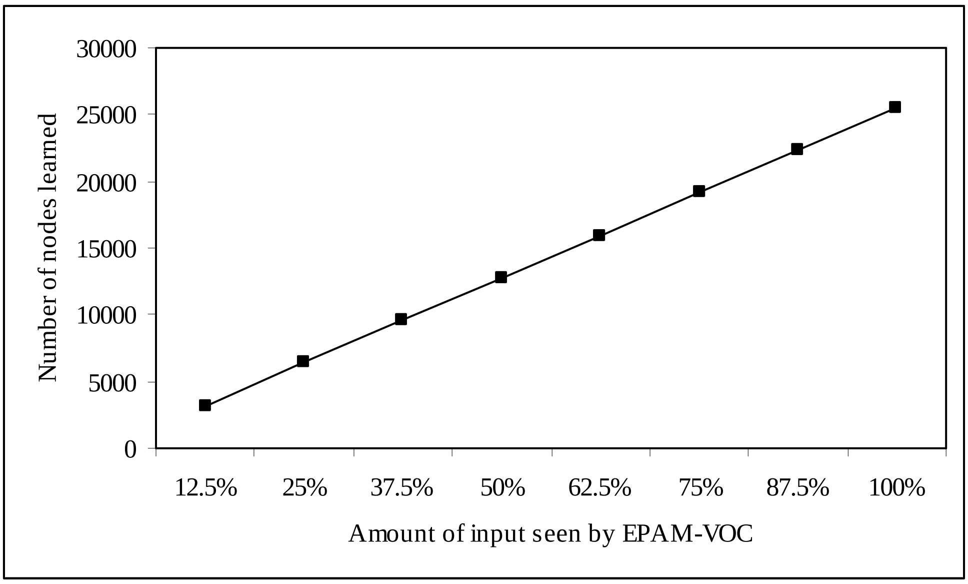

2000
1800
1600
1400
1200
1000
800
600
400
200
0
Time to match chunk (ms)
1
2
3
Syllables in non-word-like nonword
25%
50%
75%
100%

Figure 7


30000
25000
20000
15000
10000
5000
0
Number of nodes learned
12.5%
25%
37.5%
50%
62.5%
75%
87.5%
100%
Amount of input seen by EPAM-VOC