RUNNING HEAD: LEARNING NEW WORDS

Jones, G., Gobet, F., & Pine, J. M. (in press). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*.

Linking working memory and long-term memory: A computational model of the learning of new words

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, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK

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

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 phonological working memory, which is seen as a critical component in the learning of new words. However, no detailed specification of the link between phonological working memory and long-term memory (LTM) has been proposed. In this paper, we present a computational model of children's vocabulary acquisition (EPAM-VOC) that specifies how phonological working memory and LTM interact. The model learns phoneme sequences, which are stored in LTM and mediate how much information can be held in working memory. 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. EPAM-VOC shows 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 provides a simple and elegant computational account of some of the key processes involved in the learning of new words: it specifies how phonological working memory and LTM interact; makes testable predictions; and suggests that developmental changes in NWR performance may reflect differences in the amount of information that has been encoded in LTM rather than developmental changes in working memory capacity.

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) (henceforth phonological working memory), which is seen as a bottleneck in the learning of new words. According to Gathercole and Baddeley, children with a high phonological working memory capacity are able to maintain more sound patterns in memory 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. In the NWR test (Gathercole, Willis, Baddeley & Emslie, 1994) children are presented with nonwords of varying lengths and asked to repeat them back as accurately as possible. 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, 1990a; Gathercole, Willis, Emslie & Baddeley, 1992). Furthermore, the central role of phonological working memory in NWR is highlighted by the pattern of performance in adults with a specific deficit in phonological working memory. These individuals have no difficulty in learning word-word pairs, but are impaired in learning word-nonword 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, they argued that phonological working memory mediated the storage of phonological knowledge in LTM (Gathercole & Baddeley, 1989). This conclusion derived further support from the work of Gathercole, Willis, Emslie and Baddeley (1991), which compared the influence of nonword length versus the influence of familiar segments within the nonwords. Whereas increases in nonword length consistently led to a decline in NWR performance, the same was not true for increases in the number of familiar segments within the nonword, suggesting a significant role for phonological working memory in novel word learning.

However, phonological working memory is not the only factor 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 is an interaction between phonological working memory and LTM in determining NWR performance. This conclusion derives support from 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 finding suggests that phonological working memory may only influence the learning of 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 working memory and LTM combine as yet (Gathercole, Willis, Baddeley & Emslie, 1994). Gathercole and colleagues hypothesise that there is a reciprocal relationship between phonological working memory and existing vocabulary knowledge (e.g. Gathercole, Hitch, Service & Martin, 1997), and together with nonword learning, the three share a highly interactive relationship (Baddeley, Gathercole & Papagno, 1998). Nonwords are represented in phonological working memory but can be supported by phonological "frames" that are constructed from existing phonological representations in LTM (Gathercole & Adams, 1993; Gathercole, Willis, Emslie & Baddeley, 1991). Frames may contain parts of stored lexical items that share phonological sequences with the nonword contained in phonological working memory. The more wordlike a nonword is, the more it will be possible to use phonological frames to boost NWR performance. The more "novel" a nonword is, the less it will be possible to use phonological frames and the more reliance will be placed on phonological working memory.

An alternative though similar view is that it is lexical structure that influences NWR performance. Metsala (1999) suggests that a child's vocabulary growth influences lexical restructuring, with words that have a dense neighbourhood requiring more restructuring than those that have a sparse neighbourhood.

Neighbourhood density is defined as the number of other words that can be formed by the substitution, addition or deletion of one phoneme in the word. Metsala (1999) found that words with dense neighbourhoods had an advantage over words with sparse neighbourhoods when performing phonological awareness tasks, supporting the view that dense neighbourhood words had been structured at a deeper level. Moreover, further regression analyses showed that phonological awareness scores contributed unique variance in vocabulary size after NWR scores had been entered into the regression, whereas there was no unique variance when NWR 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 well-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, which would explain why performance differences in wordlike versus non-wordlike nonwords are more pronounced in younger children. Bowey (1996) argues for developmental changes in phonological processing ability whereby phonological representations become more elaborated as vocabulary size 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 new words are learned or how new words are stored in LTM and phonological working memory. Furthermore, there is no explanation of how the representations in LTM interact with those in phonological working memory.

The goal of this paper is to fill 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 four key phenomena in the NWR data. Not only is the model consistent with the explanations of the link between long-term and phonological working memory 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 in language learning, LTM is more likely to be the driving force behind the learning of new words.

We chose EPAM/CHREST as our computational architecture, as it has been used to simulate several language-related phenomena both in adults and children including phenomena in verbal learning and early grammatical development (Gobet, Lane, Croker, Cheng, Jones, Oliver, & Pine, 2001). EPAM/CHREST's individual components and mechanisms have been well validated in prior simulations, as have the node-link structures and several of the time parameters used by the architecture. EPAM/CHREST also offers a natural way of combining working memory and LTM, which opens up the possibility of integrating across time-based (e.g. Baddeley & Hitch, 1974) and chunk-based (e.g. Miller, 1956; Simon & Gilmartin,1973) approaches to working memory capacity.

Our model, which we call EPAM-VOC, simulates the acquisition of vocabulary by the construction of a network, where each node encodes a "chunk" of knowledge, in our case a sequence of phonemes. Development is simulated by the growth of the network, so that increasingly longer phonological sequences are encoded. The model also has a short-term phonological working memory, where a limited number of chunks can be stored. The exact capacity is dictated by time-based limitations, as the information held in phonological working memory is subject to decay within two seconds. Although the number of chunks that can be held in phonological working memory is limited, the amount of information, counted as number of phonemes, increases with learning, as chunks encode increasingly longer sequences of phonemes. As we demonstrate below, the interaction between the acquisition of chunks in LTM and the limitations of phonological working memory enables the model to account for the key NWR phenomena that have been uncovered in the literature.

The layout of the remainder of the paper is as follows. First, we summarise the existing NWR findings, together with a summary of existing models of NWR performance. Second, we provide a description of EPAM-VOC. Third, we present a new experiment on NWR performance, which fills a gap in the current literature: because existing studies do not use the same nonwords across ages, a developmental account of the model cannot be compared to the same datasets. Fourth, we show 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, we provide a general discussion of the findings of our experiment and our simulations.

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 higher for 2-syllable nonwords than for 3-syllable nonwords, and for 3-syllable nonwords than for 4-syllable nonwords. Second, children's repetition accuracy gets better with age. For example, Gathercole and Adams (1994) found that 5 year olds' NWR performance was 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 NWR exist, although neither was created with the intention of accounting for the key phenomena listed above. Hartley and Houghton (1996) describe 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 NWR. 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 makes similar types of error to those made by children and adults.

Brown and Hulme (1995, 1996) describe 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 subject 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 subject 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 increase in strength more than non-wordlike nonwords.

Long nonwords decay more quickly than short nonwords. The model therefore simulates the fact that children's 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 are met by either model. Furthermore, neither model explains how phonological knowledge is actually acquired through exposure to naturalistic stimuli.

Further models of short-term memory exist although they were not created with the purpose of simulating NWR performance. Brown, Preece and Hulme (2000) describe OSCAR, a model of serial order effects that is able to simulate a wide range of serial order phenomena such as item similarity and grouping effects. Page and Norris' (1998) primacy model simulates word length and list length effects using decay, and phonological similarity effects by the inclusion of a second stage of processing where similar items in the to-be-remembered list may cause phonological confusions. Burgess and Hitch's (1999) network model also simulates word length and list length effects, together with the effects of articulatory suppression and phonological similarity. The model is able to learn the sounds of words and their pronunciations by strengthening the connection between the phonemic input and the representation of the word in the network. As we saw earlier with the two models of NWR performance, none of these models are able to explain how phonological knowledge is acquired and how new words are learned. When simulations require long-term knowledge (e.g. when to-be-remembered words become confused with phonologically similar words in the lexicon) this information is added rather than learned. Even when the sounds of words are learned, as in the Burgess and Hitch (1999) model, the model itself already includes the nodes that represent the words.

In summary, several models of NWR have shed important light on the mechanisms in play. However, none of the models reviewed are able simultaneously to (a) detail how phonological knowledge is learned and how new words can be formed, (b) explain how long-term phonological knowledge interacts with phonological working memory, and (c) account for the key phenomena we have described. We now present in detail a computational model that satisfies all these desiderata.

A new computational model of nonword repetition: EPAM-VOC

EPAM (Elementary Perceiver And Memorizer, Feigenbaum & Simon, 1984) and its variants constitute a computational architecture that have been used to model human performance in various psychological domains, such as perception, learning, and memory 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 several phenomena in grammatical development (Freudenthal, Pine & Gobet, 2006, in press; Freudenthal, Pine, Aguado-Orea & Gobet, in press; 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 different domains. This body of research enables us to present a model that has very few ad hoc assumptions.

EPAM progressively builds a discrimination network of knowledge by distinguishing between features of 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.

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. EPAM provides a simple mechanism by which the sequence of phonemes can be represented in a hierarchical fashion that preserves their order. As such, the resulting discrimination network becomes a long-term memory of phoneme sequences. Preliminary versions of the model have been described in Jones, Gobet and Pine (2000b, 2005). The model in the current paper extends these preliminary versions by simplifying the learning mechanisms and taking into account the role of the input and the roles of encoding and articulation processes on NWR performance. This section will first describe how EPAM-VOC builds a discrimination network of phoneme sequences, and second, how phonological working memory will be simulated and linked to the discrimination network.

Learning phoneme sequences 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.[1] We make the simplifying assumption that, at the beginning of the simulations, EPAM-VOC has knowledge of the phonemes used in English (this assumption has support in the vocabulary acquisition literature, e.g. Bailey & Plunkett, 2002).

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 current 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 current node and the process is repeated. When a node is reached where no further traversing can be done (i.e. the next phoneme does not exist in the links below the current node, or the node has no links below it), learning occurs by adding the next phoneme in the input sequence as a link and node below the current 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 node, plus the new phoneme just added. Sequence learning, where increasingly large "chunks" of phonemes are acquired, is very similar to discrimination in traditional EPAM networks.

As stated earlier, at the beginning of the simulations EPAM-VOC is assumed to know the individual phonemes of the English language. These are stored as nodes in

[1] The simplifications include not using the familiarisation mechanism and the time parameters related to learning.

the first level below the root node. When EPAM-VOC receives an input (a sequence of phonemes), new nodes and links are created. With sequence learning, the information at nodes becomes sequences of phonemes, which in some cases correspond to lexical items (e.g. specific words) rather than just individual sounds (i.e. phonemes).

Let us consider an example of the network learning the utterance “What?”. Utterances are converted into their phonemic representation using the CMU Lexicon database (available at http://www.speech.cs.cmu.edu/cgi-bin/cmudict). This database was used because it enables automatic conversion of utterances to phoneme sequences, containing mappings of words to phonemic equivalents for over 120,000 words. For example, the utterance “What?” converts 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 as given in the database (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. The node with the link “W” is taken, and EPAM-VOC now moves on to the remainder of the input (“AH1 T”), trying to match the first part of the remaining input (“AH1”) by examining the links below the current node. Since the “W” node does not have any links below it, no further matching can take place. At this point, 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 node, so EPAM-VOC reverts back to the root node and moves on to the remainder of the input ("T"). This part of the input can be matched below the root node by taking the "T" link, but as there is no further input, no further learning takes place.

Using "W AH1 T" as input a second time, EPAM-VOC is able to match the first part of the input ("W"). The next part of the input is then examined ("AH1"), and because this exists as a link below the "W" node, the "W AH1" node becomes the current 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 and so a new node and link "T" can be made below the current node. Thus, after two successive inputs of the sequence "W AH1 T", the whole word is learned as a phoneme sequence, and the network is as shown in Figure 1. At this point, the model could produce the word "What". It should be noted that in EPAM-VOC, all information that is in the network is available for production. For children, there is a gap between comprehension and production (e.g. Clark & Hecht, 1983).[2]

Insert figure1 about here

---

[2] A possible way of differentiating between comprehension and production would be to distinguish, as does EPAM (Feigenbaum & Simon, 1984), between creating nodes in the network (mechanism of discrimination) and adding information to a given node (mechanism of familiarisation). Thus, information can be recognized as long as it is sorted to a node in the network, without necessarily assuming that the model can produce it, which would require elaboration of the information held at this node.

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. 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/CHREST 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 working memory and linking it to the discrimination network

EPAM-VOC now requires a specification of phonological working memory and a mechanism by which phonological working memory interacts with EPAM-VOC's discrimination network. When detailed (e.g. Gathercole & Baddeley, 1990b), phonological working memory is synonymous with the phonological loop component of the working memory model (Baddeley & Hitch, 1974). The phonological loop has a decay based phonological store which allows items to remain in the store for 2,000 ms (Baddeley, Thomson & Buchanan, 1975). EPAM-VOC therefore has a time-limited phonological working memory 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 assume rehearsal in order to model memory span (e.g. Brown & Hulme, 1995). EPAM-VOC therefore does not use a 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 phonological working memory is a simple time-based store, in-line with current findings regarding rehearsal in young children.

Having described the model's phonological working memory and LTM (i.e. the discrimination network of nodes and links), 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 phonological working memory links to LTM and how learning modulates this link. Within EPAM-VOC, it is relatively easy to specify how phoneme sequences in LTM interact with phonological working memory. When phonemes are input to EPAM-VOC, they are matched to those that are stored as nodes in the discrimination network; for any phoneme sequences that can be matched in LTM, a pointer to the relevant node is placed in phonological working memory.[3] That is, input sounds are

[3] There are several plausible biological explanations for the notion of a pointer. For example, short-term memory neurons in the prefrontal cortex may fire in synchrony with neurons in posterior areas of the brain, and the number of pointers that can be

not necessarily stored individually in phonological working memory, 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 phonological working memory 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 phonological working memory. 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. The first parameter (400 ms) refers to the time to match a node in LTM, create a pointer to it in phonological working memory, and "unpack" the information related to the first phoneme of this node. The second parameter (30 ms) refers to the time needed to unpack each subsequent phoneme in phonological working memory.

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 phonological working memory 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

---

held in short-term memory is a function of the number of distinct frequencies available (e.g. Ruchkin, Grafman, Cameron, & Berndt, 2003).

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). The other phonemes are stored individually 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 phonological working memory, 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 phonological working memory. 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 phonological working memory. 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 into 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 largely consistent with that suggested by Gathercole and colleagues. EPAM-VOC learns 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. Phonological sequences can therefore be seen as phonological frames in Gathercole and Adams (1993) terms.

The two accounts diverge in their explanation of how long-term phonological knowledge influences information in phonological working memory. For Gathercole and colleagues, the amount of information held in phonological working memory does not necessarily increase with increases in the number of phonological frames. Rather, phonological frames can be used to improve the quality of the encoding of items in phonological working memory at the point of retrieval, a process known as redintegration (Gathercole, 2006). For EPAM-VOC, the amount of information that can be held in phonological working memory varies as a function of the number and length of phonological sequences held in LTM. The reliance on phonological working memory as a mediator of verbal learning therefore depends on EPAM-VOC's existing phonological knowledge, 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 density. EPAM-VOC learns more detail for words with dense neighbourhoods relative to words with sparse neighbourhoods. Dense neighbourhood words by definition have many other words that differ only by a single phoneme, whereas sparse neighbourhood words do not. All other things being equal, this means that EPAM-VOC learns more about dense neighbourhood words because similar phoneme sequences are more likely to occur in the input. For example, compare the dense neighbourhood word make (which has neighbours such as take and rake) with the sparse 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 is 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 phonological working memory suggest that phonological working memory mediates NWR performance – it is a bottleneck in language learning (e.g. Gathercole, 2006). 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 one of the issues that 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 at different ages purely in terms of differences in the amount of phonological knowledge 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 in its learning, and between older children and the model at a later stage in its learning. Unfortunately, NWR studies have tended to use different sets of stimuli (Gathercole, 1995), making comparison difficult. Furthermore, existing studies have carried out NWR 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 problem 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 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 NWR 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 (as was done by Gathercole, Willis, Emslie & Baddeley, 1991). 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 (replicating the methodology of Gathercole & Baddeley, 1989). The randomised order was consistent for all children.

Nine different coloured blocks of equal size were used for a 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. After seeing each sequence of blocks the children were given all 9 blocks and asked to repeat the sequence of blocks they had just been shown. 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.

Procedure

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 to repeat back immediately exactly what they had 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 three sequences of coloured blocks (starting at length two). If two were repeated back correctly, then the length was increased by one and the process began again. 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 NWR 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 ($F(2,236)=260.52$, $Mse=116.93$, $p<.001$); for both ages and nonword types, longer nonwords were more difficult to accurately repeat. 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 NWR scores and vocabulary size ($r(58)=.49$, $p<.001$). While the correlation between NWR 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 NWR scores and vocabulary size ($r(58)=.76$, $p<.001$).

Partial correlations were also performed on the children's data. If the theory instantiated in EPAM-VOC is correct, long-term phonological knowledge should influence both NWR scores and vocabulary scores but not span scores. There should therefore be a higher correlation between NWR scores and vocabulary scores when span is partialled out than there is between span scores and vocabulary scores when NWR score is partialled out. This was, in fact, the case ($r(117)=.54$, $p<.001$ for NWR and vocabulary with span partialled out versus $r(118)=.43$, $p<.001$ for span and vocabulary with NWR partialled out).

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 NWR 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 acoustic characteristics of the one-syllable nonwords they used (Gathercole & Baddeley, 1989). For example, thip and bift showed very poor repetition performance because of the presence of fricative/affricative features (Gathercole, Willis, Emslie & Baddeley, 1991).

The correlational data are also consistent with previous findings, where significant correlations have been found between NWR 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 NWR results and the results of the correlational analysis show a high degree of consistency with previous studies of NWR, 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 in presenting each nonword as input and checking whether the model could represent the nonword within the 2,000 ms time capacity. However, children's NWR performance is clearly error prone, whereas EPAM-VOC, as described so far, has no way of producing errors, other than by being unable to represent the whole of the relevant nonword within the 2,000 ms time limitation. This means that without some additional mechanism for generating errors, EPAM-VOC would be incapable of producing errors on one-syllable nonwords since such items have a maximum of three phonemes and so would fit easily into phonological working memory – 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). To make EPAM-VOC more psychologically plausible, we introduced an error-producing mechanism according to which an incorrect link could be taken probabilistically during traversal of the network. This allows EPAM-VOC to produce repetition errors even when all phonemes can fit into phonological working memory.

After the model had 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 traversing three 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 nonwords in children, the error rates for two- and three-syllable nonwords are the product of the model's dynamics. In fact, as will be shown later, the model matches the children's performance on two- and three-syllable nonwords better than it matches their performance on one-syllable nonwords.

The probability of producing an error was reduced as more input was given to the model (see Table 2), because it was assumed that as children get older, they become more adept at encoding and articulating the sounds they receive. The probability of making an incorrect traversal was reduced at a linear rate of 1% at each point in the model's learning. That is, no attempt was made to 'fit' the error probability to the error rates of older children. At the end of the simulations, the probability of making a traversal error was .04, corresponding to a 12% error rate for single-syllable nonwords (the probability of making a correct sequence with three nodes is .96*.96*.96=.88).

<u>The input regime</u>

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 words were selected from the CMU Lexicon database. Pairs of words were used because the average number of phonemes in an utterance (across all mothers) was 12.03 whereas the average number of phonemes in a word from the CMU Lexicon database was 6.36. That is, pairs of words were used in order to maintain a similar number of phonemes in the sequences used as input.

The relative ratio of maternal utterances and pairs of random words from the lexicon were gradually altered to reflect increasing variation in input as the child grows older. Initially, the first 25% of the maternal input was given to EPAM-VOC, and thereafter progressively 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 maternal utterances to pairs of lexicon words. In terms of input, EPAM-VOC was presented with the same number of utterances that appeared in each 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 of the simulations, 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 NWR test was carried out, there would be a .09 probability of traversing down an incorrect link. Note that, in the example, every tenth utterance was replaced, meaning that the simulations for each child used the same subset of maternal input (i.e. the same utterances were replaced) but different sets of random word pairs (i.e. the pairs of words selected to replace them) differed.

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 in detail how its performance on the NWR task evolves over time.

For all simulations, input was converted into a sequence of phonemes using the CMU Lexicon database. This database cross-references words with their phonemic form. All of the phonemes used in the database map onto the standard phoneme set for American English. The phonemic input did not distinguish word boundaries, so no word segmentation had been performed on the input that was fed to the model.

Simulations of the data

A total of 120 simulations were carried out (ten for each of the sets of maternal utterances). Ten simulations per set of utterances were used in order to produce reliable 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. NWR results were averaged across the 120 simulations.

To compare EPAM-VOC with 2-3 year old children's NWR performance, an NWR test was performed after the model had seen 25% of the input (i.e. when only maternal utterances had been seen as input). To compare EPAM-VOC with 4-5 year old children, an NWR test was performed 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. There was also a significant main effect of nonword type ($F(1,238)=63.30$, $Mse=76.54$, $p<.001$), wordlike nonwords being repeated back more easily than non-wordlike nonwords. There was also a significant main effect of nonword length ($F(2,476)=310.98$, $Mse=314.86$, $p<.001$) with longer nonwords being repeated less accurately than shorter nonwords. There was an interaction between stage of learning and nonword type ($F(1,238)=18.20$, $Mse=22.00$, $p<.001$), between nonword type and nonword length ($F(2,476)=35.32$, $Mse=34.95$, $p<.001$), and between stage of learning and nonword length ($F(2,476)=42.48$, $Mse=43.01$, $p<.001$). There was also a significant three-way interaction ($F(2,476)=6.40$, $Mse=6.34$, $p<.01$).

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 for the children, there was a highly significant correlation ($r(10) = .91$, $p < .001$; 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. More specifically, the results show that EPAM-VOC meets three of the four key criteria outlined earlier in the paper. First, it simulates the finding that NWR performance declines as nonword length increases. Second, it simulates the finding that repetition performance improves as learning proceeds. Third, it simulates the finding that wordlike nonwords are repeated more accurately than non-wordlike nonwords.

However, although our new data provided a solid base on which to test the model, our new experiment did not include single and clustered consonant nonwords. It therefore provides no data on which to assess the model's ability to meet the fourth criterion (i.e. to simulate the finding that NWR performance is better for single consonant than for clustered consonant nonwords). In order to show that EPAM-VOC also meets 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 performed after the model had seen 75% of the input, and to compare to five year olds, a NWR test was performed after the model had seen 100% of the input. These input figures are consistent 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). Most importantly, these results show 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 above, 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 a maximum of two 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. Testing the error predictions of the model in more detail would require more detailed data on children's NWR errors than are currently available.

Summary of the simulations

EPAM-VOC provided a very good fit to the new data from the experiment presented here, and the model also showed a similar pattern of performance to the 4 and 5 year old children studied by Gathercole and Baddeley (1989), although the fit was not as close 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 rather low repetition accuracy for four-syllable nonwords. This suggests that EPAM-VOC had not seen enough (or varied enough) 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 build such an input set is likely to result in only a crude approximation. For example, the maternal utterances used as input only contained 3,046 different words on average, so that even when boosted with words from the CMU lexicon, the input sets used in the simulations were unlikely to capture the diversity of input that 4 and 5 year old children actually receive. The model thus provides a good fit to the existing data on children's NWR performance based on what would seem to be a reasonable, but not perfect, approximation of the distribution of phonological information in 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 phonological working memory until the model contains a large amount of phonological knowledge, 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 in 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 is provided by plotting the number of nodes that are learned 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 acquired as EPAM-VOC receives more input – and this means that EPAM-VOC is more able to fit longer nonwords within the time limit of phonological working memory.

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 phonological working memory than single consonant nonwords.

NWR performance is better for wordlike than non-wordlike nonwords

One possible explanation for this phenomenon comes from the fact that there is a slight difference in the phonemic length of wordlike and non-wordlike nonwords; this is because non-wordlike nonwords tend to have clustered consonants. In our experiment, wordlike nonwords had on average 5.00 phonemes, compared to 5.67 phonemes for non-wordlike nonwords. However, this difference in itself is unlikely to be sufficient to produce such striking performance differences on 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 phonological working memory. We can check this by subjecting the model's performance to the same ANOVA reported previously, but using the time to match nonwords as the dependent measure rather than NWR scores. This analysis 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 phonological working memory. It is thus clear that wordlike nonwords can be represented using fewer nodes than non-wordlike nonwords, which is why these nonwords are more likely to fit within phonological working memory.

General discussion

In the last decades, short-term memory capacity has been measured in two ways. Starting with Miller (1956), one group of researchers has proposed that 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, focusing on Baddeley and Hitch's (1974) model of working memory, has 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 LTM, and (b) specifies the relation between working memory and LTM.

We identified four criteria that any viable model of NWR 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 working memory 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; see also the data presented in this paper). 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 the 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 specifies how phonological working memory interacts with existing LTM phonological knowledge. Second, the simulations illustrate how differences in performance at different ages may not require explanations based around capacity differences – rather, the explanation can be 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 hierarchy of phoneme sequences in order to increase the amount of information that can be held in phonological working memory. 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 phonological working memory, 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 phonological working memory. Precisely how phonological working memory interacts with LTM has never been defined before in computational terms.

While more precise and quantitative than current views of how phonological working memory and LTM interact, EPAM-VOC's account is still consistent with them. 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 LTM. 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 learned 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 new word 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 NWR. 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 with dense neighbourhoods require more restructuring than words with sparse neighbourhoods, and thus there is more lexical structure surrounding dense 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 NWR 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 developmental variations in capacity. Gathercole, Hitch, Service and Martin (1997) suggested that the capacity of phonological working memory 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. While individual differences in capacity exist (e.g. Baddeley, Gathercole &

Papagno, 1998), the results presented here suggest that developmental differences in capacity may not be 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 the NWR task may suffer from the same problem as the digit span task.

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 working memory in much the same way as EPAM-VOC's network of phonological knowledge facilitates the amount of input that can be processed within its phonological working memory.

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 of 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 phonological working memory 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 learned 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.

Parameters used by the model

The simulations included several parameters that affect the model's ability to perform NWR, and it is worth discussing them in turn. First, an error probability was set based on 2-3 year old children's one-syllable NWR errors. The error probability was decreased at a linear rate at each stage of the model's learning. Using a linear reduction in error probability means that no serious attempt has been made to select error probabilities that would 'fit' the later error rates of older children (in fact a better alternative would be for the error rates in the model to be an emergent property based on, for example, how often phoneme sequences are accessed in the network). Nevertheless the results presented would benefit from a detailed analysis of NWR performance at varying levels of error probability. Decreases in probability would be expected to improve accuracy for short nonwords to a greater extent than long nonwords, because the model already has difficulty in representing long nonwords in phonological working memory. Increases in error probability would be expected to show more of a decline for long nonwords because these contain more phonemes and are therefore more prone to error when making traversals in the network.

Second, the actual input seen by the model influences what the model learns. The input given to EPAM-VOC was intended to approximate that of 2-5 year old children, with the assumption that 2-3 year olds' linguistic input can be estimated from the utterances produced by their primary caregiver and 4-5 year olds' linguistic input can be estimated from a mixture of utterances from the primary caregiver and random words from a lexicon. From the results shown, these would seem reasonably reliable estimates.

Third, comparisons were made to the children's data at various stages of the model's learning that reflected the age of the child (e.g. comparing 2-3 year olds NWR performance with the model after 25% of the input had been seen). The stages chosen were 25% (2-3 year olds), 75% (4 year olds), 87.5% (4-5 year olds), and 100% (5 year olds). As with the error probability parameter, it is clear that no attempt has been made to select stages of learning that would optimally 'fit' the children's data, but again it would be beneficial to examine the pattern of change in NWR performance when NWR tests are performed at a variety of different stages of the model's learning. NWR performance would be expected to be poor at earlier stages of learning in the model because EPAM-VOC would not have built up enough phoneme sequences in LTM, but performance would be expected to improve as the model is presented with more input.

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, phonological working memory). At later stages of the model's learning, emphasis is placed on LTM. 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 in 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 learned. An obvious strength 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 may arise from relative experience with components of the task at hand.

EPAM-VOC is obviously only a first attempt at simulating the learning of new words. There are clearly areas where the model requires further development. 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. Furthermore, the relatively simple way in which phonological working memory is implemented means that when nonwords are unable to fit in the time-limited store, they are cut-off such that only the beginning part of the nonword is repeated. By contrast, children tend to maintain nonword length even though constituent syllables may be incorrect (Marton & Schwartz, 2003). The model could be improved by considering further findings in the vocabulary acquisition and memory literature, and considering other computational models in this area (e.g. Burgess & Hitch, 1992).

Nevertheless, the model presented here provides a simple and elegant computational account of some of the key processes involved in the learning of new words and is able to simulate the NWR findings surprisingly well. In addition, EPAM-VOC reconciles time-based and chunk-based approaches to memory capacity. In doing so, it provides well-specified mechanisms on the relation between working memory and LTM, 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.

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

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.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. Psychological Review, 107, 127-181.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. Psychological Review, 106, 551-581.

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.

Clark, E. V., & Hecht, B. F. (1983). Comprehension, production, and language acquisition. Annual Review of Psychology, 34, 325-349.

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.

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.

Freudenthal, D., Pine, J. M., & Gobet, F. (in press). Understanding the developmental dynamics of subject omission: The role of processing limitations in learning. Journal of Child Language.

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

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

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

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

Gathercole, S. E. & 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. (1990a). 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., & Baddeley, A. D. (1990b). Phonological memory deficits in language disordered children: Is there a causal connection? Journal of Memory and Language, 29, 336-360.

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., 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 learning. 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.

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. Journal of Speech, Language, and Hearing Research, 46, 1138-1153.

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.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. Psychological Review, 105, 761-781.

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.

Ruchkin, D. S., Grafman, J., Cameron, K. & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. Behavioral and Brain Sciences 26, 709-777.

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. Note that although only five individual phonemes are illustrated below the root node, the model knows all phoneme primitives.

Figure 2. Repetition accuracy for EPAM-VOC after 25% of the input, plotted against 2-3 year old children. Standard deviations for each data point can be found in Table 1.

Figure 3. Repetition accuracy for EPAM-VOC after 87.5% of the input, plotted against 4-5 year old children. Standard deviations for each data point can be found in Table 1.

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). Note that no error bars are given for the child data because Gathercole and Baddeley (1989) do not specify standard deviations.

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). Note that no error bars are given for the child data because Gathercole and Baddeley (1989) do not specify standard deviations.

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

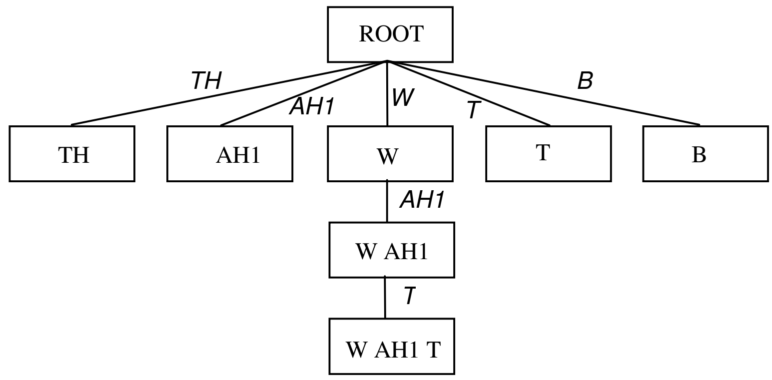


Figure 2

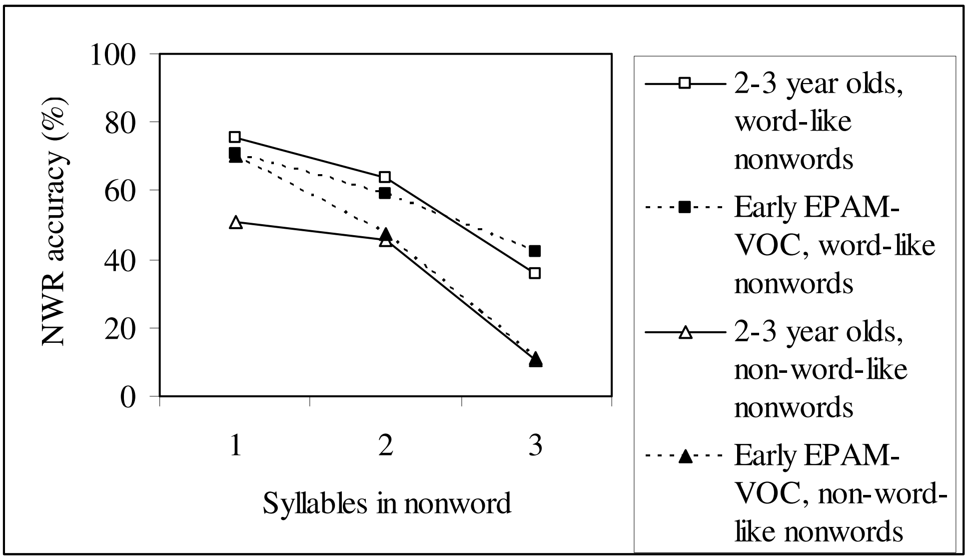

100
80
60
40
20
0
NWR accuracy (%)
1
2
3
Syllables in nonword
2-3 year olds, word-like nonwords
Early EPAM-VOC, word-like nonwords
2-3 year olds, non-word-like nonwords
Early EPAM-VOC, non-word-like nonwords

Figure 3

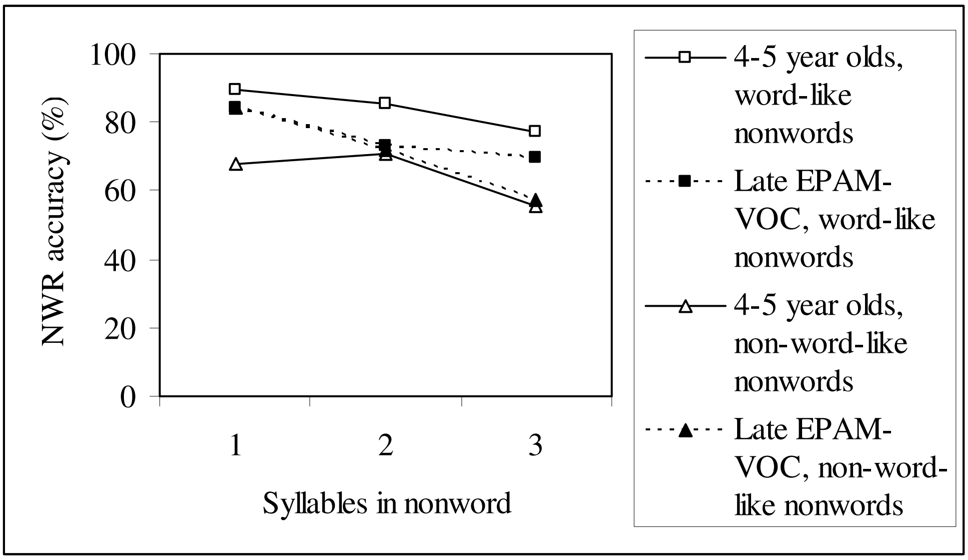

100
80
60
40
20
0
NWR accuracy (%)
1
2
3
Syllables in nonword
4-5 year olds, word-like nonwords
Late EPAM-VOC, word-like nonwords
4-5 year olds, non-word-like nonwords
Late EPAM-VOC, non-word-like nonwords

Figure 4

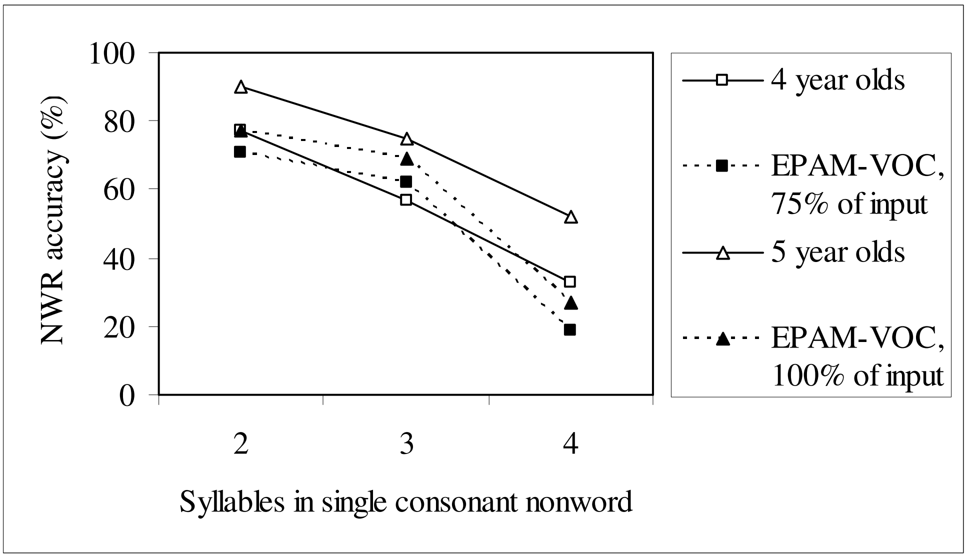

100
80
60
40
20
0
NWR accuracy (%)
2
3
4
Syllables in single consonant nonword
4 year olds
EPAM-VOC, 75% of input
5 year olds
EPAM-VOC, 100% of input

Figure 5

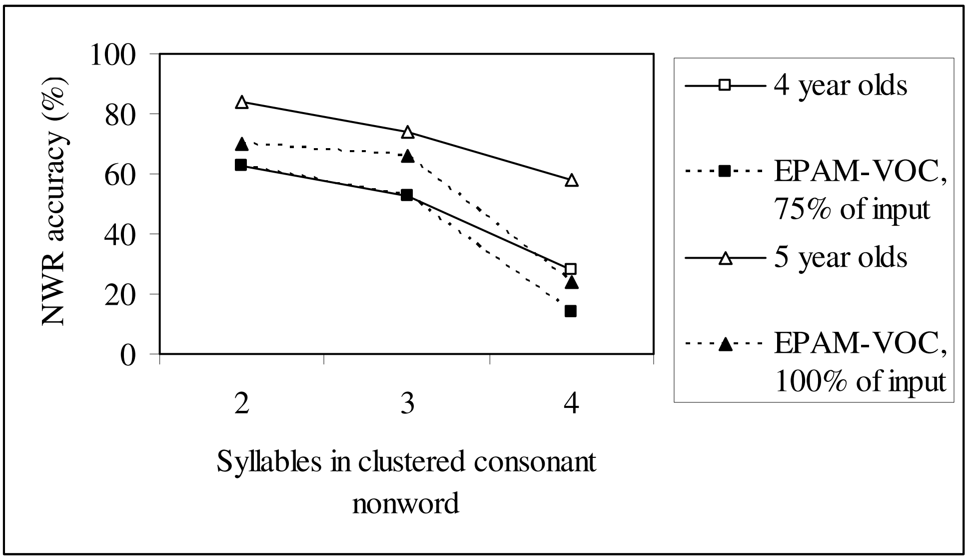
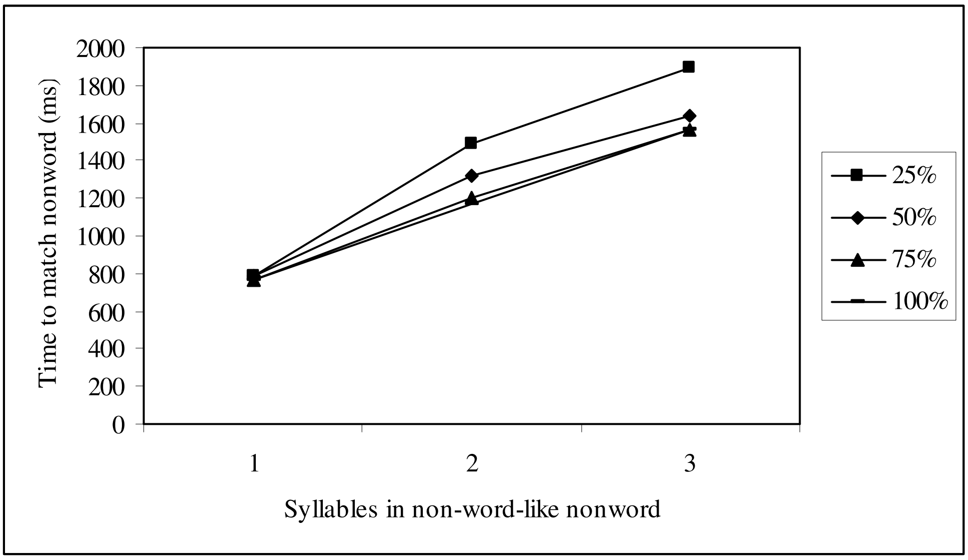
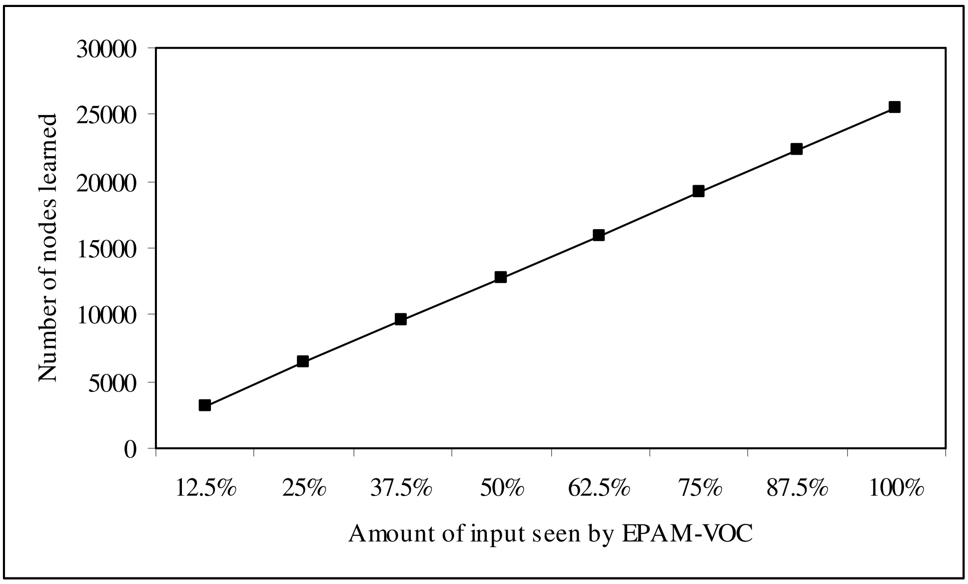

100
80
60
40
20
0
NWR accuracy (%)
2
3
4
Syllables in clustered consonant nonword
4 year olds
EPAM-VOC, 75% of input
5 year olds
EPAM-VOC, 100% of input

Figure 6



Time to match nonword (ms)
2000
1800
1600
1400
1200
1000
800
600
400
200
0
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