26

Working memory and SLA

John N. Williams

Introduction

The term *working memory* (WM) refers to a temporary storage system that lies at the core of complex cognition. Everyday tasks such as remembering a phone number, mental arithmetic, or playing chess, place varying levels of demand on the WM system. Remembering an unfamiliar phone number long enough to dial it involves storing the number in short-term memory, and perhaps repeating it silently to oneself. Multiplying 26 by 7 in your head involves storing the digits produced by multiplying 6 and 7 (4 and 2) while multiplying 2 and 7, adding the carried 4 and retrieving the stored 2 to give the final answer of 182. Skilled chess playing involves strategic planning of a sequence of moves whilst retaining multiple piece configurations in short-term memory. These tasks not only require short term storage of information, but also a high degree of control of that information in order to maintain partial results whilst carrying out further processing, or to resist interference between similar items of information. Thus, WM can be regarded as a system that is used for the temporary maintenance of task-relevant information whilst performing cognitive tasks.

As one would expect, individuals vary a good deal in their ability to carry out such complex cognitive tasks, or even, as it turns out, relatively simple ones like remembering sequences of digits. This gives rise to the notion of working memory capacity as an important dimension of individual difference (see Skehan, this volume). The question that will concern us here is whether variation in WM capacity (WMC) is related to variation in first and second language processing and learning ability. The answer to this question could not only prove useful in using tests of WM as a predictor of language abilities, but also could illuminate the nature of the processing and learning mechanisms themselves.

Historical discussion

There are two essential components to the WM system. The first is concerned with the temporary storage of information; that is, short-term memory. The second is concerned with the control of that information, as required to carry out complex tasks, and the component responsible for this is variously referred to as the *Central Executive* (Baddeley, 2007) or *executive attention* (Kane *et al.*, 2007). The first model of WM was proposed by Baddeley and Hitch (see 1974 for a review). They postulated two domain-specific short-term memory systems—the phonological loop for the storage of phonological information (as used when remembering a phone number) and the visuo-spatial scratch pad for visual and spatial information (as used when remembering chess configurations). It is the first of these that will concern us most here of course. Information in these systems was assumed to decay rapidly - in the case of information in the phonological store after about two seconds. In order to maintain information for longer it is necessary to refresh its activation through a rehearsal process. Hence the phonological loop contains a passive phonological store, which will be referred to here as *phonological short term memory* (PSTM), and a subvocal rehearsal process (a metaphorical "loop"). The central executive was conceived as a mechanism for controlling and managing information in the short-term stores. In Baddeley's early work, where the focus was primarily on the functioning of the phonological loop, little attention was paid to the central executive beyond its role in controlling rehearsal processes. The phonological loop was explored using simple random digit and word list recall tasks in which it is natural to employ rehearsal strategies. It was found that recall is poorer for similar sounding items (the "phonemic similarity effect") reflecting interference between representations in the phonological store. Recall was also found to be affected by factors that affect rehearsal rate, such as the length of the items (the "word length effect"), and a person's articulation rate, effects that disappear when rehearsal is prevented by concurrent articulatory suppression tasks (Baddeley, 2007).

Other researchers have developed the original Baddeley and Hitch (1974) conception of WM to make it more relevant to complex cognition in general, focusing directly on the role of the central executive in maintaining task-relevant information (Cowan, 1999; Engle, Kane and Tuholski, 1999). It is this latter function that is crucially important in the mental arithmetic and chess examples considered earlier. Here the problem is not just the storage of phonological or visuo-spatial information but also the maintenance of specific representations in the face of potential interference from similar items in memory, or stimuli from the environment. Kane *et al.* (2007) provide other everyday examples such as searching for your car in a car park that you use frequently or when driving on the opposite side of the road in a foreign country. In the first case memories of other spaces you have previously used in the car park interfere with your search (a case of "pro-active interference"). In the second, the problem is to manage the competition between automatic responses and the novel ones required by the current context. Thus, the executive component of the WM system is assumed to be involved even in tasks that do not have an obvious short-term storage component. Indeed, there is good evidence that this component of WM is related to general fluid intelligence, as measured by non-verbal figural reasoning tasks (Engle, *et al.*, 1999).

At first sight it may seem that very different notions of WM abound in the literature. Yet these are differences of emphasis rather than overall conception. Most researchers agree that WM is a multi-component system comprising domain-specific storage systems and a domain-general executive component. Whilst the focus of much of Baddeley and colleagues has been on the storage components, the focus of Engle and colleagues has been on executive functions. As we shall see, both of these perspectives can make different, yet complementary, contributions to our understanding of SLA.

This chapter is organised in two parts, the first dealing with PSTM specifically and the second with the operation of the WM system as a whole. It is easiest to understand the various different aspects of WM by considering how they are measured, and so a discussion of elicitation measures and empirical verification will precede discussion of core issues. The chapter will end with a consideration of applications and future directions.

Data and common elicitation measures: Phonological short-term memory and SLA

Traditional measures of verbal short-term memory involve immediate repetition of sequences of varying numbers of random digits, words, or non-words in the order of presentation. For example, the maximum number of random digits that a person can accurately and reliably repeat back in the correct order is known as their *digit span*. Typically, items are presented at a rate of one per second providing the opportunity to cumulatively rehearse the sequence. Hence the correspondence between digit span and factors that affect rehearsal rate, such as word length. Following the Baddeley and Hitch WM model, performance on this task is predicted to be a function of the decay rate of information in the phonological store (PSTM) as well as the efficiency of the sub-vocal rehearsal process, as determined by the person's articulation rate. Performance is also dependent upon whether the person chooses to adopt a rehearsal strategy. These factors introduce ambiguity into the interpretation of digit span measures. For example, the fact that children have smaller digit spans than adults could be because of faster decay, slower articulation rate, or failure to use rehearsal. Indeed rehearsal strategies typically do not develop until around age 7 (Baddeley *et al.*, 1998).

The *non-word repetition task* is often used as an alternative, and arguably more direct, measure of verbal short-term memory, and may be particularly suitable for use with young children (Gathercole *et al.*, 1994). Participants are simply required to immediately repeat back individual nonsense words of varying lengths (e.g. *ballop, doppelate, empliforvent*). The longer the non-word, the more likely it is that the beginning will have been lost from PSTM before it has ended. The larger a person's PSTM capacity, the longer the non-word that they will be able to repeat successfully. Note that, as operationalised by this task, PSTM "capacity" refers to the duration, rather than the amount, of information in PSTM. The test measures how long phonological information persists in the absence of rehearsal.

An important issue that impacts upon the design and interpretation of the non-word repetition test, however, is that performance is affected by lexical knowledge, such as similarity to known words (Gathercole, 1995). For example, English-French bilinguals are equally good at repeating non-words conforming to English and French phonotactics, but monolinguals are worse for non-words conforming to French phonotactics (Thorn and Gathercole, 1999). It has also been argued that non-word repetition only partially reflects short-term memory capacity but also a variety of phonological processing skills such as speech perception, phonological encoding, and assembly of a phonological representation for articulation (Bowey, 2006). Thus, as in the case of immediate serial recall, the test cannot unambiguously be interpreted as a simple, direct, measure of PSTM capacity.

Given the differences between immediate serial recall and non-word repetition tasks it is hard to see how they can both be regarded as tests of the same underlying construct of PSTM. Digit span uses highly familiar items and invites rehearsal strategies, non-word repetition uses unfamiliar items and rehearsal is not possible. Yet performance on the two tasks has been found to correlate (Gathercole *et al.*, 1992; Gupta, 2003). The correlation between the tasks suggests that they share a common component of short-term storage ability, possibly the ability to retain serial order. The problem in a digit span task is not only remembering which digits occurred in the list, but also the order in which they occurred. Likewise, when remembering a non-word the problem is remembering the order of the segments, not just the segments themselves. It has been suggested that short-term memory for item and order information should be regarded as distinct aspects of short-term memory (Gupta, 2003).

Short-term memory tests have been developed that are more sensitive to order than item information. In the *serial non-word recognition* task varying length sequences of single syllable CVC non-words are constructed (e.g. *mel, guk, vip*) which are either immediately followed by the same sequence, or by a sequence containing a reversal of two adjacent items (e.g. *guk, mel, vip*). The task is simply to indicate whether the two sequences are the same or different. Performance on this task is hardly affected by whether the items are words or non-words (Gathercole *et al.*, 2001), or whether the words are from a person's L1 or L2 (Thorn *et al.*, 2002) suggesting that it is principally sensitive to memory for sequence, rather than item, information.

Another way of measuring order information independently of item information is by using the *serial order reconstruction task* (Majerus *et al.*, 2008). This is a variant of the digit span task in which instead of recalling the sequence orally the participant is provided with cards corresponding to the digits used in the sequence and is required simply to arrange them into the correct order. The digits used in each sequence are known in advance because, for example, a five-item sequence uses the digits 1 to 5, and a six-digit sequence the digits 1 to 6. Thus there is no requirement to remember the items, only the order in which they occurred.

Core issues: PSTM and language learning

Vocabulary learning

An early indication that PSTM is implicated in language learning came from the short-term memory patient PV (Baddeley *et al.*, 1998). Despite being severely impaired on digit and word span tasks she was able to learn arbitrary associations between pairs of known words in her native Italian. However, she was unable to learn unknown L2 (Russian) words paired with L1 (Italian) translations. This suggested that PSTM plays a specific role in the formation of long-term phonological representations. Subsequently, in a study of normally developing children Gathercole *et al.* (1992) found that both digit span and non-word repetition were correlated with vocabulary size, the relationship being strongest for non-word repetition in four and five year-olds. A similar relationship has been found in adult SLA (Service and Kohonen, 1995) and in laboratory studies of foreign vocabulary learning (Service and Craik, 1993; Speciale *et al.*, 2004; Williams and Lovatt, 2003).

Of course, given that non-word repetition is itself related to vocabulary knowledge there is a potential circularity in using it as a predictor of vocabulary development. The similarity of a novel word to known words could affect both ease of repetition, and also the ease of storage in LTM. However, PV showed impaired novel word learning even though she had unimpaired access to her native vocabulary. Furthermore, digit span has been found to be related to vocabulary learning, even though digit recall makes minimal demands on lexical knowledge (Atkins and Baddeley, 1998). Gupta (2003) even found that digit span was a better predictor than non-word repetition of learning non-word names for pictures of imaginary animals. These are compelling results given that digit span would appear to be a very different task from vocabulary learning.

Gutpa (2003) suggested that the connection between digit span and vocabulary learning is largely due to the serial order component of PSTM. Majerus *et al.* (2008) tested this hypothesis by examining the efficiency of learning pairs of English and French-derived non-words (a laboratory simulation of vocabulary learning) by English learners of French at varying levels of proficiency. Performance on the serial order reconstruction task correlated highly with word-non-word learning, as did French proficiency, and regression analyses showed that these factors made independent contributions to word-non-word learning. Majerus *et al.* assume that the proficiency measure provides an indication of the extent of support for vocabulary learning from the existing French lexicon. They propose that proficiency relates to item retention in PSTM, whereas the serial order reconstruction task relates to retention of sequence information, and that both factors contribute to vocabulary learning. Evidence for an independent contribution of the efficiency of sequence learning to vocabulary learning is also provided by Speciale *et al.* (2004).

Although there is good evidence for a relationship between PSTM and vocabulary learning it should be noted that this has not always been evident in studies conducted outside the laboratory, particularly when learners already have some familiarity with the language. Gathercole *et al.* (1992) found that non-word repetition is predictive of subsequent vocabulary size only at the very earliest ages tested (between age 4 and 5). Masoura and Gathercole (2005) studied Greek children who had been studying English for three years and found that although there was a strong relationship between non-word repetition (using English-derived non-words) and current English vocabulary size, non-word repetition ability did not predict the ability to learn new words. Likewise, in a study of learning of English by French children in a five-month intensive programme, French and O'Brien (2008) found that non-word repetition ability as measured at the start of the programme only explained 3.5 per cent of the variance in vocabulary development. These children had already received 100 hours of English instruction prior to the study. Gathercole (2006) suggests that PSTM might be most important at the initial stages of language learning. As the size of the vocabulary develops word learning is affected more by the ease of integrating words into the existing lexicon through words of similar sound and meaning, reflecting LTM contributions to learning of new information. Thus, even individuals with quite poor PSTM ability can eventually develop normal-sized vocabularies through applying these other learning strategies. This does not mean that PSTM is not a relevant consideration in naturalistic language learning, but it is one of many factors that determines vocabulary growth.

Grammar learning

Emergentist and connectionist perspectives on language acquisition imply a close relationship between acquisition of vocabulary and grammar. Learning words involves learning sequences of phonemes, and learning grammar involves learning sequences of morphemes (Ellis, 1996). Formulaic memorised chunks can provide the data for further analysis, leading to the abstraction of rules (Myles *et al.*, 1999). So to the extent that vocabulary acquisition is dependent upon PSTM, and to the extent that learning sequences of morphemes is like learning sequences of phonemes, then PSTM would be predicted to be related to grammar learning.

Indeed there is good evidence for a close relationship between vocabulary size and grammatical development in normal and abnormal populations (Bates and Goodman, 1997). Larger vocabularies provide a richer database from which morphosyntactic regularities can be extracted. And there is experimental evidence that when the operation of the phonological loop is disrupted through articulatory suppression, both the acquisition of vocabulary and grammatical rules are severely affected (Ellis and Sinclair, 1996). But these observations do not necessarily imply a direct relationship between PSTM and grammar learning. For instance, as more vocabulary items are learned, segmentation of grammatical morphemes becomes easier, and these segmented morphemes could trigger learning processes that are not necessarily dependent upon PSTM at all. What is the evidence that PSTM influences grammar learning independently of its effect on vocabulary learning?

One approach to answering this question is to examine grammar learning in situations where knowledge of vocabulary is already well-developed. This was the case in Robinson's (1997) study of various components of aptitude in relation to acquisition of grammatical rules by intermediate-level learners of English. He found that memory ability was related to rule learning, but only in the instructed and rule search training groups. No such relationship was found for the implicit (memory task) and incidental (comprehension task) groups. This study provides evidence of a connection between PSTM and grammar learning that is not mediated by vocabulary learning, at least when explicit learning strategies are being used.

In a study by Williams (1999) participants with no prior knowledge of Italian first learned the vocabulary items to be used in the experiment and then performed a memory task on sentences exemplifying a range of inflectional morphemes and agreement patterns. Rule learning was assessed by a surprise grammaticality judgment task on sentences containing known words in novel combinations. Rule learning was correlated with the speed with which the vocabulary had been learned in the pre-training phase, and also the accuracy of sentence recall over the first few training trials. Assuming that these measures reflect PSTM ability, this experiment provides indirect evidence for a connection between PSTM and grammar learning.

More direct evidence comes from a study by Williams and Lovatt (2003, Experiment 2) that focused on the induction of arbitrary noun classes. In order to learn the target system participants had to remember which articles had occurred with which nouns in the input; that is, they had to infer the underlying noun class distinction from distributional information alone (there were no semantic or phonological cues to class membership). Participants first learned the form and meaning of all of the articles and nouns to be used in the experiment as isolated units. They were then exposed to grammatical combinations of articles and nouns in a memory task. Learning of the noun class distinction was assessed by cyclic generalisation tests on novel combinations of the articles and nouns used in training. Both PSTM (as measured by non-word span) and prior knowledge of gender languages made independent contributions to learning. Presumably PSTM affected the ability to remember article-noun combinations during the training task, and these formed the basis for inducing the underlying noun class distinction.

Given that the above studies are all laboratory-based it is important to obtain evidence for the PSTM-grammar learning link from more natural situations. French and O'Brien (2008) examined the relationship between PSTM, as measured by non-word repetition, and grammar learning in French children in an intensive five-month English immersion program. Non-word repetition at the start of the programme was correlated with grammatical knowledge at the end, a relationship that held even when initial grammatical knowledge, nonverbal intelligence, L2 contact, and, crucially, vocabulary knowledge at the end of the programme had all been entered into the regression model. French and O'Brien suggest that, although there were improvements in vocabulary knowledge over the program, the independent contribution of PSTM to grammar learning reflects learning sequences of morphemes, as opposed to the morphemes themselves.

Thus, there is good evidence from both laboratory- and classroom-based studies for a contribution of PSTM to grammar learning. This appears to result from a relationship between PSTM function and the ability to store morpheme sequences in LTM, even when the forms of those morphemes are already familiar.

Data and common elicitation measures: Working memory and SLA

Following the psychological construct of WM outlined earlier, tests of WMC need to measure a person's ability to both retain and manage information in short-term memory in the face of potential interference from other cognitive tasks. Daneman and Carpenter (1980) developed a reading span task (RST) in which participants read increasingly longer sequences of unrelated sentences and then have to recall the final word of each sentence in order. A person's reading span is the maximum number of sentence-final words that can be reliably recalled. This is clearly more complex than PSTM tasks because participants are required to maintain increasing numbers of sentence-final words in STM whilst simultaneously processing irrelevant sentences. Performance on the reading span task varies considerably between individuals (between 2 and 6 in college students according to Daneman and Carpenter, 1980). Hence working memory capacity, as operationalised by this task, could be an important dimension of individual differences.

However, span tasks of this type can only be used to investigate individual differences if they are reliable. Waters and Caplan (1996) found that the test re-test reliability of the Daneman and Carpenter test was poor, as was the consistency of classification of participants into span groups, even over short time periods. One reason for this may be that participants vary in the resources they devote to the secondary, sentence reading, task. Waters and Caplan found that a variant that required participants to judge each sentence for plausibility, rather than just read it aloud, had higher reliability and stability. In other variants, each sentence is followed by an unrelated word that has to be remembered (Engle, Tuholski *et al.*, 1999), or by a single letter (Kane *et al.*, 2004). This prevents people from using memory for sentence gist to aid recall of sentence-final words.

The Daneman and Carpenter RST was based on a domain-specific view of WM. In this view, the best way to measure the WM resources used in reading comprehension is with a task that involves memorisation of verbal material whilst reading sentences. In contrast, for Engle and colleagues (Engle *et al.*, 1999; Engle *et al.*, 1999) the executive component of WM is domain-general and it is only necessary that the task combines a memory component with a demanding processing component. They therefore developed span tasks that do not necessarily involve sentence comprehension. In their operation span task (Turner and Engle, 1989) the sentences of the reading span task were replaced with simple equations that the participants had to verify (e.g. 9/3–2 = 1, or 9/3–2 = 6) followed by single words that had to be recalled. In counting span tasks (e.g. Engle *et al.*, 1999) participants have to count the number of target items (e.g. dark circles) in a series of displays containing distracters (dark blue squares and light blue circles). At the end of the series they have to recall the number of target items that were in each display. Conway *et al.* (2005) suggest that the simplicity of the processing component in counting span tasks makes them particularly suitable for a range of populations, including children and brain-damaged patients. The fact that the task is language independent would also appear to make it particularly suitable for use in second language research. See Conway *et al.* (2005) for a discussion of other methodological issues in relation to the administration and scoring of WM tests.

Correlational studies support the contention that WM tasks are primarily sensitive to domain-general executive function and only secondarily sensitive to domain-specific storage abilities, the converse being the case for STM tasks (Caplan *et al.*, 2007). For example, WM tasks predict around 50 per cent unique variance in general fluid intelligence (Kane *et al.*, 2005) compared to only 18 per cent for PSTM tasks (Conway *et al.*, 2002). Verbal and spatial WM tasks share 70–85 per cent of variance, but verbal and spatial STM tasks only 40 per cent (Kane *et al.*, 2004).

Core issues: Working memory and language learning

First language ability

Daneman and Carpenter (1980) argued that language comprehension is a prime example of a complex cognitive activity that involves temporary maintenance of task-relevant information during ongoing processing. They found that reading spans of college students were highly correlated with their reading comprehension test scores, with these correlations lying between 0.5 and 0.6 in various experiments. When specific reading abilities were considered, for example the ability to answer a factual question about a passage, or to determine the referent of a pronoun, the correlation with reading span reached 0.9. (see Daneman and Merikle, 1996 for a meta-analysis). In studies of young children WMC has also been shown to be related to anaphor resolution (Oakhill and Yuill, 1986), and to the ability to resolve anomalies when the critical information is distant in the text (Yuill *et al.*, 1989). On the other hand, attempts to relate WMC to on-line syntactic processing have led to largely negative results, leading to the hypothesis that syntactic processing, at least in native speakers, is modularized and does not draw on general verbal working memory capacity (Caplan *et al.*, 2007). Correlations between WMC and tests of language comprehension are attributed to higher-level "post-interpretive" comprehension processes (e.g. anaphor resolution). However, some recent studies have revealed an influence of WMC on syntactic processing when very distant elements have to be integrated (Havik *et al.*, 2009; Roberts *et al.*, 2007), suggesting that WMC limitations are only evident when storage and integration costs (Gibson, 1998) of on-line processing are very high.

Second language ability

A large number of studies have found correlations between L1 and L2 RST performance, with coefficients of.39 (Harrington and Sawyer, 1992), 0.48 (Berquist, 1997), 0.61 (Juffs, 2004), 0.84 (Osaka and Osaka, 1992), 0.85 (Osaka *et al.*, 1993). This is hardly surprising, given the assumed domain-generality of the executive component of WM, although what accounts for the variability in strength of association is not clear. For example, both Harrington and Sawyer (1992) and Osaka and Osaka (1992) studied advanced Japanese learners of English, and yet found very different strengths of correlation.

Reading span has also been reported to be lower in the L2 than in the L1 (Berquist, 1997; Havik *et al.*, 2009; Osaka *et al.*, 1993; Walter, 2004) whereas in some studies not significantly different (Harrington and Sawyer, 1992; Osaka and Osaka, 1992). A problem here is to equate the difficulty of RST tests in different languages since when the syntax of the languages differs radically, the sentence-final words will have very different properties (Harrington and Sawer, 1992; Osaka *et al.*, 1993). An obvious solution for obtaining better comparisons would be to require recall of words or letters that follow each sentence.

It seems reasonable to assume that where differences between L1 and L2 reading spans are obtained this is partly attributable to level of L2 proficiency. Indeed, Harrington and Sawyer (1992) found that L2 RST correlated with TOEFL (Test of English as a Second Language) scores at $r = 0.54$, and Berquist (1997) found a correlation with TOEIC (Test of English for International Communication) scores at $r = 0.41$. Walter (2004) found that lower intermediate French learners of English had lower English RST scores than upper intermediate learners, and this held even when differences in L1 RST performance were taken into account. Reduced L2 RST performance is to be expected on a single resource view (Just and Carpenter, 1992), since as the processing drains of sentence processing increase the WM resources available for storage of sentence-final words decreases. It could also reflect the influence of differences in domain-specific storage capacity on WM function (Engle, Kane *et al.*, 1999). For example, if L2 representations are of "low quality" (Perfetti, 2007) they may be more subject to interference. Performance on an L2 RST would be reduced, whilst the common domain-general executive attention ability would ensure that L2 and L1 performance is also correlated.

With regard to specific tests of L2 comprehension, Walter (2004) found correlations between an L2 RST and performance on L2 summary completion and anaphor resolution tasks. These correlations were stronger for lower intermediate learners, and not significant at all for the upper intermediate learners on the anaphor resolution task. This pattern is to be expected if WMC constraints are most apparent when the processing system is most challenged (Just and Carpenter, 1992). Kormos and Safar (2008) found a correlation of 0.55 ($p < 0.01$) between WMC (as measure by L1 reverse digit span) and performance on the Cambridge First Certificate Exam after the first year on an intensive English instruction program. Correlations with WM were also found over the Reading, Listening, Speaking, and Use of English components of the exam.

With regard to on-line syntactic processing, one might expect that increased difficulty, or reduced automaticity, in L2 might make general verbal WMC limitations more evident than in L1. However, studies have either shown no effect of WMC on L2 sentence processing (Juffs, 2004, 2005), or an effect in L1 but not L2 (Felser and Roberts, 2007). One exception is Havik *et al.* (2009) who showed that both L1 and L2 sentence processing are similarly constrained by WMC. In general, though, there is surprisingly little evidence for WMC constraints on on-line L2 syntactic processing. Thus, the mass of evidence for a general relationship between WMC and L2 language skills would seem to be largely attributable to post-interpretive processes beyond the sentence level.

Reasoning, category learning, and artificial grammar learning

There are good reasons to expect that WM should be related to learning in general, but only when learning processes involve intentional control; that is, broadly speaking, explicit, as opposed to implicit learning. First we shall consider evidence from reasoning, category learning, and artificial grammar learning that suggests that this should be the case, before considering studies of SLA in the following section.

Reasoning. WM, as well as IQ, are strongly related to what Evans (2003) dubs explicit "System 2" conscious abstract reasoning and hypothetical thinking, but not implicit "System 1" reasoning, as supported by phylogenetically older and domain-specific associative systems (a distinction that aligns with Reber's distinction between explicit and implicit learning systems (Reber (1993)). Interestingly, the relationship between conscious reasoning and IQ is particularly strong when semantic information has to be ignored in favour of logical form, as for example when verifying the logical correctness of the syllogism "All mammals can walk, whales are mammals, therefore whales can walk" (see Stanovich and West, 2000, for a review). It is reasonable to conclude that this relationship reflects the executive attention component of WM.

Category learning. DeCaro *et al.* (2009; DeCaro *et al.*, 2008) demonstrate a link between WM, as measured by reading and operation span tasks, and inductive learning of rule-based categorisations, where the categories can be stated as easily verbalisable rules (e.g. if the stimulus embedded within the object is a circle then the object belongs to category A). High WMC individuals were also found to be superior on more complex categorisations that involve integrating information over a number of dimensions, although in this case it is not clear whether the effect was due to implicit learning or complex explicit knowledge (DeCaro *et al.*, 2008). Low WM individuals appear to persist with simple learning strategies when faced with complex problems (see also Beilock and DeCaro, 2007). In the case of the category learning problem studied in DeCaro *et al.* (2008, 2009) this could produce above-chance responding, and even superior performance to high WM individuals if a lax criterion for success was adopted. However, only high WM individuals were able to learn the categories according to stricter criteria (DeCaro *et al.*, 2009). Hence the apparent success, or even superiority, of low WM individuals in complex learning tasks can be more apparent than real.

Artificial grammar learning. There is good evidence that learning is related to IQ under explicit, rule discovery, but not implicit, memorise, instructions (Gebauer and Mackintosh, 2007; Reber, Walkenfeld, and Hernstadt, 1991). In so far as tests of general fluid intelligence are assumed to tap the executive component of the WM system (Engle, Kane *et al.*, 1999) then in this sense it can be argued that WM is related to explicit, but not implicit, learning. However, Robinson (2002, 2005) found that low WM was actually associated with more success in artificial grammar learning under instructions to memorise. It is possible that this reflects spontaneous explicit learning in high WM individuals, which may have actually suppressed performance (see Reber, 1976, for evidence that rule search can lead to worse performance than memorisation in artificial grammar learning). This is another reason why low WM might lead to superior performance.

Second language learning

In light of the above findings it certainly makes sense to predict a role for WM in SLA, especially when learning processes towards the explicit end of the spectrum are involved. It is therefore clearly important to consider this question in relation to different learning conditions, as far as the available data allow.

Explicit instruction. Roehr (2008) hypothesises that working memory should be important in learning from explicit instruction because of the requirement to retain metalinguistic information in memory whilst simultaneously producing and comprehending language, and because explicit rule-based processing should place demands on the central executive. Some evidence for this is provided by Ando*et al.* (1992, described in Mackey *et al.* 2002) who found that L1 WM correlates with L2 learning after 20 hours of explicit form-focused instruction, although only on a two-month delayed post-test.

Task-based learning. Here there is an emphasis on interaction and the role of interactional feedback in prompting learners to briefly focus on form in a meaningful context. WM might be important for rapidly switching attention between form and meaning, for making comparisons between one's own output and a corrective recast, and in general for permitting greater degrees of noticing in communicative situations (Doughty, 2001; Robinson, 2001). A study by Mackey *et al.* (2002) provides initial support for these ideas. They examined the development of question forms in Japanese learners of English. Teaching was through communicative tasks that involved interactional feedback followed up by stimulated recall activities to evaluate noticing. High working memory was associated with more noticing of relevant forms (although the effect was marginally significant over the small sample tested). In subsequent studies WMC has been found to be positively associated with the probability that a learner will modify their utterance in response to feedback whilst engaging in an interactive task with a native speaker (Mackey *et al.*, 2010; Sagarra, 2007). Thus there is good evidence that higher WMC enables greater noticing of feedback and subsequent modification of output. In terms of actual learning gains, Mackey *et al.* (2002) found evidence for more development amongst the high WM group at the delayed post-test (although this analysis is only based on a total of 7 participants). In a computer-based task requiring typed cloze responses Sagarra (2007) found that providing the correct answer as feedback (i.e. recasts) produced high learning gains as measured by immediate and delayed post-tests, and that learning gains correlated with WMC. Although this was not a face-to-face communicative task, the results do suggest that WMC is involved in learning from feedback.

Intentional induction. Brooks, Kempe, and Sionov (2006) and Kempe and Brooks (2008) examined learning of a Russian inflectional system in which there were associations between noun suffixes and semantic (case) and noun class (grammatical gender) cues. Participants performed comprehension and production tasks with feedback that would have encouraged them to search for regularities. Learning was assessed by performance on old items and generalisation to new items. In Brooks *et al.* (2006) both non-verbal fluid intelligence and WM made independent contribution to performance on old items, but only IQ made an independent contribution to new items. Following Engle, Tuholski *et al.* (1999) they argue that tests of general fluid intelligence tap the executive attention component of the WM system, and that this is related to learning the underlying rules of the target system because of the requirement to direct attention to distributional patterns during the training task. The relationship between WM and performance on old items reflected the storage component of the WM system. Similar results were obtained by Kempe and Brooks (2008) for a non-transparent system in which there were no phonological cues to noun class membership.

Incidental learning. Robinson (2002, 2005) examined learning of some syntactic rules of Samoan under "incidental" training conditions that simply involved reading sentences for meaning and answering comprehension questions. There were no effects of WM on grammaticality judgement post-tests that used visually presented sentences. However, there were positive correlations with WM on a listening version of the grammaticality judgement task and a production task, especially in the delayed post-tests. It is not clear whether these correlations reflect the effect of WM on learning, or whether they reflect the effect of WM on test performance. Once again, note that WM correlations were most evident in the post-tests. Also note that according to the questionnaire data the participants had high levels of awareness of two of the three target rules, and a high proportion of them said that they had searched for rules during the training task. Thus, it appears that if WM had affected learning it was through the mechanism of intentional induction.

There is mounting evidence for a relationship between WM and second language learning, and this has come from studies that seem to have involved explicit learning, as one would expect from work done in the related areas of reasoning, category learning, and artificial grammar learning.

Applications

Skehan (this volume) suggests that the broad areas of application of individual differences research relate to selection, counselling, remediation, and instructional modification. In particular, ultimately one would like to see tests of WM being used as a means of directing students to instructional programs that are tailored to their cognitive abilities. For example, students with low PSTM might require a relatively low rate of vocabulary acquisition, with plenty of opportunities for recycling. Students with low WM might struggle with tasks that rely on intentional inductive learning where they are required to discover patterns for themselves, or may have difficulty learning from interactive tasks where rapid switching of attention between form and meaning is necessary. However, it must be born in mind that before WM tests are used in this way we need far more evidence of the relationship between different aspects of WM and specific learning processes (see *Future Directions* below).

Another approach to application lies not in modifying instruction, but in modifying cognitive abilities themselves. There is now evidence that WM can be improved through training (Holmes, Gathercole, and Dunning, 2009; Klingberg, 2010). These effects are not confined to specific tasks, but have been shown to generalise from, for example, spatial to verbal memory, and in particular to tasks that involve the executive component of WM, such as Stroop tests. In fact, Holmes *et al.* (2009) found effects of training on WM, but not PSTM, tasks, suggesting that WM training improves the executive, rather than the storage, component of the WM system. Note, however, that most of this work has been carried out on children and special populations (e.g. children with attention-deficit hyperactivity disorder (ADHD)). Also, the fact of being a fluent bilingual appears to be associated with enhanced executive functions, presumably because of the increased control demands of managing two languages (Bialystok, 2009). Apart from providing further evidence that WMC is modifiable by experience, this also illustrates a potential application of WM research in encouraging bilingualism and bilingual education.

Future directions

Clearly it is a difficult task to specify exactly how a process as multi-faceted as SLA is related to a psychological construct that is as complex as WM, but some possible ways of homing in more precisely on the basis of their relationship are as follows:

(1) Pay more attention to the nature of the learning targets. For example, in relation to PSTM and grammar learning, do the target rules have to depend on phonological distinctions

(e.g. agreement patterns, distinguishing articles that otherwise have the same meaning)? If PSTM is to have a more general role, say, in learning word order regularities, then retention of sequential information will be critical, and the relevant item information will have to be represented at the more abstract level of meaning or grammatical categories. In general, if we are to understand the precise role of PSTM in grammar learning more attention needs to be paid to the nature of the regularities under investigation, specifically the kind of information required for rule induction, and also to the potential importance of both item and order information as dissociable aspects of PSTM function.

(2) Pay more attention to learning tasks, contrasting in particular implicit and explicit learning modes. We need to be more precise about what learners are actually doing during learning tasks in order to relate learning success to specific cognitive functions. This may be achieved through a combination of more constrained learning tasks and think-aloud and stimulated recall procedures (as used for example by Mackey *et al.*, 2002).

(3) Compare more or less complex (or verbalisable) learning targets, as in the category learning studies of DeCaro *et al.* (2008, 2009) and the inflection learning study of Kempe and Brooks (2008). This will potentially reveal how different learning processes along the explicit-implicit dimension are related to WM.

(4) Pay more attention to specific WM components. The reading span task involves both storage and executive functions it is difficult to relate performance to precise learning processes. More direct, non-linguistic, tests of executive function might be more useful, such as counting and calculation span tasks. The use of non-verbal IQ is a step in this direction (Brooks *et al.*, 2006; Kempe and Brooks, 2008), but this appeals to an assumed relationship between IQ and WM.

Conclusion

There are clearly individual differences in WMC, both in relation to PSTM and executive function. We have seen that this variability impacts upon second language *learning* in various ways. PSTM ability affects the efficiency of learning novel word forms, and the retention of sequences of forms. The latter plausibly contributes to grammatical development through processes of analysis. Where analysis depends upon conscious, intentional, explicit learning processes the executive component of WM appears to be implicated. Thus, the construct of WM is undoubtedly an important component of the notion of language learning aptitude, at least in the context of these specific kinds of learning processes. As we gain a greater understanding of how WM is implicated in different learning processes we will be able to have more confidence in using tests of WMC as predictors of language learning success in specific learning contexts. At the same time we will see how individual differences in WMC can be used as a tool for prising apart different aspects of the language learning process.

References

Ando, J., Fukunaga, N., Kurahashi, J., Suto, T., Nakano, T., and Kage, M. (1992). A comparative study on two EFL teaching methods: The communicative and the grammatical approach. *Japanese Journal of Educational Psychology*, *40*, 247–256.

Atkins, P. W. B. and Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*, 537–552.

Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

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

Baddeley, A. D. and Hitch, G. (1974). Working memory. In G. Bower, (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Bates, E. and Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*, 507–584.

Beilock, S. L. and DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 983–998.

Berquist, B. (1997). Individual differences in working memory span and L2 proficiency: Capacity or processing efficiency? In A. Sorace, C. Heycock, and R. Shillcock (Eds.), *Proceedings of the GALA '97 Conference on Language Acquisition* (pp. 468–473). Edinburgh: Human Communication Research Centre, University of Edinburgh.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*, 3–11.

Bowey, J. A. (2006). Clarifying the phonological processing account of nonword repetition. *Applied Psycholinguistics*, *27*, 548–552.

Brooks, P. J., Kempe, V., and Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology. *Applied Psycholinguistics*, *27*, 185–209.

Caplan, D., Waters, G., and Dede, G. (2007). Specialized verbal working memory for language comprehension. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Eds.), *Variation in working memory* (pp. 272–302). Oxford: Oxford University Press.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., and Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*, 769–786.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.

Daneman, M. and Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Daneman, M. and Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*, 422–433.

DeCaro, M. S., Carlson, K. D., Thomas, R. D., and Beilock, S. L. (2009). When and how less is more: Reply to Tharp and Pickering. *Cognition*, *111*, 415–421.

DeCaro, M. S., Thomas, R. D., and Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition*, *107*, 284–294.

Doughty, C. J. (2001). Instructed SLA: Constraints, compensation and enhancement. In C. Doughty and M. Long (Eds.), *Handbook of second language acquisition*. Oxford: Blackwell.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, *18*, 91–126.

Ellis, N. C. and Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology*, *49A*, 234–250.

Engle, R. W., Kane, M. J., and Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefontal cortex. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454–459.

Felser, C. and Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, *23*, 9–36.

French, L. M. and O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*, 463–487.

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

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

Gathercole, S. E., Pickering, S. J., Hall, M., and Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, *54*, 1–30.

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

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

Gebauer, G. F. and Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 34–54.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *56*, 1213–1236.

Harrington, M. and Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25–38.

Havik, E., Roberts, L., van Hout, R., Schreuder, R., and Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, *59*, 73–112.

Holmes, J., Gathercole, S. E., and Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*, F9–F15.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, *102*, 199–225.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, *21*, 121–151.

Just, M. A. and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., and Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). Oxford: Oxford University Press.

Kane, M. J., Hambrick, D. Z., and Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., and Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology-General*, *133*, 189–217.

Kempe, V. and Brooks, P. J. (2008). Second Language Learning of Complex Inflectional Systems. *Language Learning*, *58*, 703–746.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*, 317–324.

Kormos, J. and Safar, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism-Language and Cognition*, *11*, 261–271.

Mackey, A., Adams, R., Stafford, C., and Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*, 501–533.

Mackey, A., Philp, J., Egi, T., Fujii, A., and Tatsumi, T. (2002). Working memory, interactional feedback and L2 development. In P. Robinson and P. Skehan (Eds.), *Individual differences and second language instruction*. Philadelphia: Benjamins.

Majerus, S., Poncelet, M., Van der Linden, M., and Weekes, B. S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. *Cognition*, *107*, 395–419.

Masoura, E. V. and 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.

Myles, F., Hooper, J., and Mitchell, R. (1999). Interrogative chunks in L2 French. *Studies in Second Language Acquisition*, *21*, 49–80.

Oakhill, J. and Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. *Language and Speech*, *29*, 25–37.

Osaka, M. and Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, *30*, 287–289.

Osaka, M., Osaka, N., and Groner, R. (1993). Language-independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society*, *31*, 117–118.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–94.

Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge*. Oxford: Oxford University Press.

Reber, A. S., Walkenfeld, F. F., and Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 888–896.

Roberts, L., Marinis, T., Felser, C., and Clahsen, H. (2007). Antecedent priming at trace positions in children's sentence processing. *Journal of Psycholinguistic Research*, *36*, 175–188.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, *47*, 45–99.

Robinson, P. (2001). Attention and memory in SLA. In C. Doughty and M. Long (Eds.), *Handbook of second language acquisition*. Oxford: Blackwell.

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on incidental second language learning: A replication and extension of Reber, Walkenfield and Hernstadt (1991). In P. Robinson and P. Skehan (Eds.), *Individual differences and second language instruction* (pp. 211–265). Philadelphia: Benjamins.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, *27*, 235–268.

Roehr, K. (2008). Linguistic and metalinguistic categories in second language learning. *Cognitive Linguistics*, *19*, 67–106.

Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 229–248). Oxford: Oxford University Press.

Service, E. and Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language*, *32*, 608–623.

Service, E. and Kohonen, V. (1995). Is the relation between phonological memory and foreign- language learning accounted for by vocabulary acquisition. *Applied Psycholinguistics*, *16*, 155–172.

Speciale, G., Ellis, N. C. and Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, *25*, 293–320.

Stanovich, K. E., and West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–665.

Thorn, A. S. C. and Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology*, *52*(A). 303–324.

Thorn, A. S. C., Gathercole, S. E., and Frankish, C. R. (2002). Language familiarity effects in short-term memory: The role of output delay and long-term knowledge. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, *55*, 1363–1383.

Turner, M. L. and Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Walter, H. C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics*, *25*, 315–339.

Waters, G. and Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, *49A*, 51–79.

Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition*, *21*, 1–48.

Williams, J. N. and Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, *53*, 67–121.

Yuill, N., Oakhill, J., and Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology*, *80*, 351–361.