Cogn Process (2008) 9:19–28
DOI 10.1007/s10339-007-0183-x

REVIEW

The role of the episodic buffer in working memory for language processing

Mary Rudner · Jerker Rönnberg

Received: 19 March 2007 / Revised: 6 June 2007 / Accepted: 3 September 2007 / Published online: 5 October 2007

Abstract A body of work has accumulated to show that the cognitive process of binding information from different mnemonic and sensory sources as well as in different linguistic modalities can be fractionated from general executive functions in working memory both functionally and neurally. This process has been defined in terms of the episodic buffer (Baddeley in Trends Cogn Sci 4(11):417–423, 2000). This paper considers behavioural, neuropsychological and neuroimaging data that elucidate the role of the episodic buffer in language processing. We argue that the episodic buffer seems to be truly multimodal in function and that while formation of unitary multidimensional representations in the episodic buffer seems to engage posterior neural networks, maintenance of such representations is supported by frontal networks. Although, the episodic buffer is not necessarily supported by executive processes and seems to be supported by different neural networks, it may operate in tandem with the central executive during effortful language processing. There is also evidence to suggest engagement of the phonological loop during buffer processing. The hippocampus seems to play a role in formation but not maintenance of representations in the episodic buffer of working memory.

Keywords Working memory · Episodic buffer · Sign language · Executive functions · Hippocampus

Working memory is often described as a mental workbench for the processing and temporary storage of information that enters consciousness, and provides a useful conceptual tool for investigating conscious processes, such as the production and understanding of language in a communicative context. Recently, evidence has been accumulating to indicate that it is theoretically interesting to consider that working memory specifically accommodates processes concerning binding of information in different codes from different sources. In this paper we review evidence in favour of isolating such a component, consider its significance in relation to language processing and discuss its neural representation.

Models of working memory

One of the most influential models of working memory is Baddeley's classic component model (Baddeley 1986; Baddeley and Hitch 1974). This model postulates two processing and storage loops and a controlling central executive, along with the more recently added integratory episodic buffer (Baddeley 2000). Whereas the loops are characterized by their specialization, verbal storage and processing in the phonological loop and visuospatial storage and processing in the visuospatial sketchpad, the theoretical notion of the episodic buffer is characterized by multidimensional storage and processing.

Another model of working memory that has been prominent in relation to language and communication is the capacity theory (Just and Carpenter 1992). Instead of postulating functionally distinct components of working memory, capacity theory suggests that cognitive capacity is limited by an available budget of activation and that, within this budget, activation can be allocated flexibly. Once all

---

M. Rudner (✉) · J. Rönnberg
The Swedish Institute for Disability Research,
Department of Behavioural Sciences and Learning,
Linköping University, Linköping, Sweden
e-mail: mary.rudner@liu.se

the available capacity has been allocated, however, any new storage or processing can be accomplished only by reducing the level of activation elsewhere.

The working memory framework for ease of language understanding (ELU) (Rönnberg 2003a; Rönnberg et al. 2007) postulates a component for Rapid, Automatic, Multimodality Binding of PHOnology (RAMBPHO) which serves to integrate multisensory, multilingual and long-term memory-based information, typically in an implicit stream. Implicit processing in RAMBPHO is mediated by the clarity of phonological representations in working memory, the speed at which they can be processed and the capacity that is available for processing. Implicit processing in RAMBPHO facilitates ELU. If one or more of these mediating factors is compromised, for example if the language signal is degraded or if processing speed or capacity are low, there is a risk of mismatch occurring between representations in working memory and long-term memory. If mismatch occurs, language understanding becomes effortful, and explicit storage and processing resources are recruited to find a key to language understanding. The ELU framework also postulates that working memory for language processing is relatively independent of language modality.

Working memory and language processing

One of the strengths of the component model is its analytical power which has provided a vehicle for a wealth of empirical work that has served to delineate its components and further our understanding of working memory. In relation to language processing, it is primarily the phonological loop that has been at the focus of attention. The phonological loop comprises a temporary phonological store in which auditory memory traces decay over a period of a few seconds, unless revived by articulatory rehearsal, and displays a number of characteristic effects including the phonological similarity effect, the word-length effect and the effect of articulatory suppression (Baddeley 2000). Where a memorized list of items has to be reproduced in the correct order, words that are phonologically similar are harder to remember accurately than words that do not sound similar and long words are harder to remember accurately than short words. These two effects are known as the phonological similarity effect and the word length effect, respectively, and together they suggest that the representation of information in the phonological loop of working memory is based on surface form characteristics of the underlying linguistic representations. Memory for word lists also deteriorates when articulation of an irrelevant sound prevents rehearsal of to-be-remembered items. This is known as the articulatory suppression effect. This phenomenon removes both the phonological similarity effect and the word-length effect, but only when words are presented visually. This is because auditory presentation bypasses the process of recoding that is thought to take place in the rehearsal loop when words are presented visually.

Baddeley (2000) has proposed that the phonological loop may have developed to support speech perception and production and the learning of new vocabulary and its pronounced reliance on serial order makes it well suited for speech-based language processing. Neuropsychological data (Vallar and Baddeley 1987) suggest that the phonological store serves as a backup system for comprehension of speech under taxing conditions, but it may be less important for straightforward communication (Baddeley 1992). This suggestion fits in with the ELU framework (Rönnberg 2003a; Rönnberg et al. 2007).

Clearly, the phonological loop of working memory underpins the processing of speech-based language but what working memory systems support the processing of signed languages?

Working memory and sign language

Signed languages are the natural languages of deaf people. Linguistically, they can be described in the same way as spoken languages but the visual and motor systems on which they rely for production and comprehension are very different. This makes them an important tool for investigating language-related cognitive processes. In a series of studies, Wilson and Emmorey (1997, 1998, 2003) have shown that working memory for sign language has an architecture similar to working memory for speech-based language as defined by the component model. In particular, there seems to be a phonological loop for sign language that is sensitive to the visual surface form of these gestural languages in much the same way that the phonological loop for speech-based language is sensitive to auditory surface form. However, loop capacity for visuospatially based lexical signs is lower than for words in both deaf and hearing native signers (Boutla et al. 2004; Rönnberg et al. 2004), suggesting that the efficiency of phonological loop processing is dependent on the temporal aspects of sound or sound-based phonological representations. On the other hand, general working memory capacity for language processing seems to be equivalent for sign and speech (Boutla et al. 2004; Rudner et al. 2007a). Neurally, working memory for sign language seems to be supported by the same structures as working memory for speech-based language (Buchsbaum et al. 2005; Rönnberg et al. 2004; Rudner et al. 2007a). However, additional structures are engaged for working memory for sign language, including the occipitotemporal region bilaterally, the superior parietal region bilaterally and the right frontal region. We have shown that these sign-language specific regions are related to the processing of individual items in working memory rather than a particular sign language mind-set, suggesting that they reflect processing of information in the individual signs (Rudner et al. 2007a). In the same study we found that binding of semantic information in signs and words engaged the right middle temporal region. Thus, the sign-specific regions may be related to the processing of non-semantic information in signs such as shape, location and movement. However, we do not know to what extent these processes are specifically linguistic or more generally cognitive. Overall, the evidence suggests that similar working memory processes support both sign- and speech-based language, but that there are some sign-specific components that may be specifically linguistic.

Working memory and long-term memory

The original component model of working memory (Baddeley and Hitch 1974) postulated a dissociation between working memory and long-term memory on the basis of neuropsychological data indicating that long-term memory could remain intact even if working memory was deficient. However, later versions of the component model have modified that standpoint (Baddeley 1986, 1996; Repovs and Baddeley 2006). Other groups have also pointed to the need for a model of working memory that takes long-term memory into account. For example, Ericsson and Kintsch (1995) proposed a model of working memory, which includes a mechanism for skilled use of storage in long-term memory. The working memory framework for ELU (Rönnberg 2003a; Rönnberg et al. 2007) emphasizes the importance of smooth access to semantic representations in long-term memory. Moscovitch (1992) put forward a neuropsychological model, in which input modules and central systems deliver their output to working memory, where it becomes conscious. The hippocampus then binds the information and encodes the resulting long-term memory trace. In the same vein, Ranganath and Blumenfeld (2005) put forward a neuropsychological model that rejects the notion of a double dissociation of long and short-term memory and proposes instead that there are separate memory stores for different types of information that retain information across both long and short delays. Baddeley (1996) postulated that the ability to relate the content of working memory to long-term memory was a function of the central executive. More recently, this function has been fractionated from the central executive to form, along with other integrative functions, the episodic buffer (Repovs and Baddeley 2006).

The episodic buffer

The episodic buffer is postulated as a separate limited capacity system within working memory that uses a multimodal code. The term “episodic” has been used to denote memories of personal happenings and doings as opposed to semantic memories that concern knowledge of the world that is independent of a person’s identity and past (Tulving 1983). In designating a working memory component, the term “episodic” indicates involvement of complex structures, or episodes, while the term “buffer” denotes that the component interfaces with other perceptual and mnemonic systems (Repovs and Baddeley 2006). The function of the episodic buffer can be described in terms of both processing and storage. As regards processing, the buffer serves as a workbench for assembling unitary multidimensional representations in different codes (visual, phonological, semantic, etc.) from different perceptual (visual, auditory, tactile, etc.) and mnemonic (episodic, semantic, etc.) sources (Baddeley 2000; Repovs and Baddeley 2006). As regards storage, the episodic buffer temporarily accommodates unitary multidimensional representations (Repovs and Baddeley 2006). Thus, the episodic buffer accommodates both the formation and maintenance of unitary multidimensional representations.

Evidence of the formation of unitary multidimensional representations in working memory was provided by the cases of two amnesic patients with severely compromised long-term memory function but preserved immediate prose recall (Baddeley and Wilson 2002). Immediate prose recall requires binding of semantic information retrieved from long-term memory with a working memory representation. Despite good immediate recall, these patients were unable to remember the prose passage after a delay. The number of words in the prose passage used for testing far exceeded normal short-term memory capacity. Thus, immediate prose recall was dependent on information chunking, which requires access to semantic representations in long-term memory. As the patients’ long-term memory was severely compromised, the inference was made that processing of words on the basis of semantic information was taking place in working memory. This could not be well explained in terms of the original three-component model (Baddeley 1986), but was in line with other findings of close links between working memory and long-term memory, which can now be interpreted in terms of the episodic buffer.

Evidence of the maintenance of unitary multidimensional representations in working memory was provided by a set of studies, showing interference between phonological and visuospatial representations in working memory (Logie et al. 2000). Visually presented words that were phonologically and visually similar (e.g. fly, cry, dry; hew, new, and few) resulted in poorer recall than words that were phonologically similar but visually distinct (e.g. guy, sigh, lie; who, blue and ewe), even when there was no articulatory suppression. This suggests that visually presented verbal sequences are stored in working memory in a visual code as well as in a phonological code, or possibly in a multidimensional code that cannot be accommodated by a unimodal loop. From a developmental perspective it has been suggested that the episodic buffer of working memory is functional in children as young as four years (Alloway et al. 2004).

The function of the episodic buffer replaces and extends the function of relating the contents of working memory to the contents of long-term memory that was originally assigned to the central executive. This suggests a close relation between the central executive and the episodic buffer, and on the basis of this it has been suggested (e.g. Baddeley 2000) that episodic buffer function may rely on central executive function in such a way that if the episodic buffer is taxed then the central executive will also be engaged. Support for this notion was provided by amnesic patients in the study by Baddeley and Wilson (2002) who showed dissociation between information binding and long-term memory processing. These patients were characterized by relatively well preserved executive functions. However, other work (Gooding et al. 2005) has shown that good executive function is not generally characteristic of densely amnesic patients, even when immediate prose recall is good. This suggests that the central executive may not necessarily support episodic buffer processing.

The episodic buffer and the central executive

The role of executive function in episodic buffer processing was explicitly addressed in a series of five experiments by Allen et al. (2006). A range of attention-demanding tasks was performed during temporary storage in working memory of bound and non-bound visual information relating to shapes and colours. In all cases, it was found that retention of bound and non-bound features was equally affected by the distracter task, suggesting that executive function plays a minor role in maintenance of bound shape and colour information. This finding may be interpreted as suggesting that maintenance of unitary multidimensional representations in the episodic buffer is not always effortful. Other work draws similar conclusions. For example, Rossi-Arnaud et al. (2006) conducted a series of experiments to investigate the impact of symmetry on a visuospatial working memory task. Results showed that while vertical symmetry aided recall, suggesting a retrieval of pattern information stored in long-term memory, processing of symmetry information did not engage executive processes. These findings are compatible with an episodic buffer that does not rely on executive processes. Similarly, Jeffries et al. (2004) conducted a series of experiments to examine the effect of an attention-demanding concurrent visual choice reaction time task on the recall of auditorily presented stories, sentences and lists of unrelated words. Unlike unrelated sentences, stories allow opportunity for chunking or integration of phonological perceptual information with semantic representations in long-term memory, which is a function of the episodic buffer. It was found that the concurrent task interfered with the processing of unrelated sentences but not with the processing of stories, suggesting that episodic buffer function is not necessarily reliant on central executive function unless loop capacity is exceeded. Thus, behavioural studies on the basis of both linguistic and nonlinguistic processes, as well as neuropsychological studies on the basis of linguistic processing provide evidence that the episodic buffer is dissociated from the central executive of working memory.

Dual processing capacity and language processing

One of the most potent predictors of performance on taxing linguistic tests is the reading span test (Daneman and Carpenter 1980), a general working memory task on the basis of capacity theory (Just and Carpenter 1992). The reading span task requires participants to read a series of short sentences and subsequently recall either the first or last word from each sentence in serial order. The test typically starts with two sentences and increases to the number of sentences at which participants are no longer able to recall all words. This number designates the subject's working memory span. Analytically, the key to the reading span test is that it involves dual processing in that it requires simultaneous processing and short-term storage. However, it has proved difficult to isolate what types of processing and storage are critically involved and how they interact (Bayliss et al. 2005). Thus, any or all of the components of working memory as defined by the component model may be tapped by the reading span task.

The reading span test predicts reading comprehension in adults (Daneman and Carpenter 1980) and primary school achievement (Gathercole et al. 2004). It is also a reliable predictor of language understanding under taxing conditions, for example, reading comprehension in children with cochlear implants (Vass et al. 2006) the ability to follow a telephone conversation after cochlear implantation in adulthood (Lyxell et al. 1998), speechreading (Rönnberg 2003a, b), visual-tactile speech recognition (Rönnberg 1993) and speech recognition in noise (Foo et al. 2007; Lunner 2003). This demonstrates that when language understanding becomes effortful because of poor phonological definition of the input, or pressure on speed of complexity of processing, the dual storage and processing capacity of working memory, as tapped by the reading span test, comes into play as predicted by the ELU framework (Rönnberg 2003a; Rönnberg et al. 2007).

The predictions of the ELU framework were tested in an experiment where hearing aid users trained for nine weeks using their own hearing aids with individually fitted frequency amplification but with adapted time constants which alter the rate of release of frequency compression. The participants' aids were set for either slow release of compression (460 ms) or fast release (40 ms). Slow release time gives quasi-linear amplification, which preserves syllable characteristics to a high degree, whereas fast release gives nonlinear amplification, which results in syllabic compression and thus a somewhat distorted sound at syllable level. Speech recognition performance with two qualitatively different kinds of noise was tested, using each of the two time constant settings both before and after training. It was found that reading span performance predicted speech recognition in both kinds of noise and with both time constant settings before training (Foo et al. 2007) and that after 9 weeks of training the correlation pattern was polarized in such a way that correlation coefficients fell for matched settings, i.e. when the participants were tested on the settings to which they had become accustomed, but rose for mismatched settings, i.e. when the participants were tested on the settings to which they had *not* become accustomed (Rudner et al. 2007b). This finding provided direct support for the ELU framework (Rönnberg 2003a; Rönnberg et al. 2007) which predicts that if mismatch occurs during the essentially implicit process of RAMBPHO, language understanding becomes effortful, and explicit storage and processing resources are recruited to find a key to language understanding.

The episodic buffer and language processing

In terms of the component model of working memory, mismatch data may further our understanding of the nature of the episodic buffer. Speech recognition involves the retrieval of lexical items from long-term memory on the basis of phonological information in the speech signal. By definition, this is a function of the episodic buffer. Speech recognition in noise involves the retrieval of lexical items from long-term memory on the basis of degraded phonological information. If, in addition to being degraded, the phonological information is distorted in relation to lexical representations, due to a difference in compression settings, we have shown that hearing aid users rely on their general working memory capacity as measured by the reading span test. We have argued that the reading span test may tap any or all of the components of working memory as defined by the component model. Thus, the increased reliance on general working memory capacity when there is a mismatch between perceptual input to the phonological loop and representations in long-term memory may reflect greater load placed on the episodic buffer. In terms of the capacity theory (Just and Carpenter 1992) and the ELU framework (Rönnberg 2003a; Rönnberg et al. 2007), this means that explicit storage and processing capacity underpin working memory in effortful language processing. In terms of the component model (Repovs and Baddeley 2006), it may mean that episodic buffer processing becomes more effortful when there is a mismatch between perceptual phonological information and stored lexical representations in long-term memory.

In order to determine what particular aspects of explicit storage and processing capacity in working memory are crucial during effortful language processing it may be useful to turn to the more analytical component model which originally postulated the phonological loop (but not the visuospatial sketchpad) and the central executive as key cognitive components in language processing; latterly the role of episodic buffer in language processing has also been highlighted.

Andersson and Lidestam (2005) examined the cognitive processing skills of the deaf speechreading expert AA and found that AA excelled on certain tests that they considered to tap central executive function, while achieving normal scores on others. The tests at which AA excelled were semantic and phonological fluency which require the retrieval of representations from long-term memory on the basis of semantic and phonological cues in working memory; thus, they tap the ability of relating the contents of working memory to long-term memory, a function which was originally assigned to the central executive but which is now considered to be accommodated by the episodic buffer (Repovs and Baddeley 2006).

Other tests administered to AA and considered by Andersson and Lidestam (2005) to tap central executive function cannot be readily redefined in terms of the episodic buffer. These tests include lexical and semantic decision-making and physical matching as well as various measures of focused attention on the basis of the Stroop task, and on these tasks AA scored in the normal range. Thus, the case of AA provides evidence on the one hand of a dissociation of episodic buffer function from the central executive and on the other hand evidence of a crucial role for the episodic buffer in language processing with a poorly defined signal.

Future work should focus on the role of the episodic buffer in the processing of audiovisual speech and sign language processing. These investigations may be usefully informed by our knowledge of the role of the episodic buffer in visuospatial processing.

The episodic buffer and visuospatial processing

Investigation of the visuospatial sketchpad of working memory has provided further evidence of the existence of the episodic buffer. The notion of the visuospatial sketchpad started life as a processing loop that dealt with visual and spatial information but was organized in a manner analogous to that of the phonological loop (Baddeley 1986; Baddeley and Hitch 1974). However, later investigation provided evidence for dissociation of visual and spatial subsystems (Klauer and Zhao 2004). Lehnert and Zimmer (2006) conducted a series of three experiments to investigate the spatial subsystem, by studying short-term memory for the location of sounds and pictures presented at varying locations in space. They hypothesized that if there are two independent spatial stores for auditory and visual representations, memory performance should be higher for location sequences in which sounds and pictures are mixed than for sequences in which they are separate. The results of all three experiments showed similar patterns of memory performance across mixed and separate sequences, suggesting that spatial information is stored independently of input modality, a notion that presupposes an integratory process such as that postulated for the episodic buffer.

Another function of the visuospatial sketchpad is the manipulation of mental imagery (Ganis et al. 2004). In an fMRI study (Rudner et al. 2005), we studied the neural correlates of manipulating auditory mental imagery using a word reversal paradigm with auditory word presentation, in which participants were instructed to imagine how the words would sound if temporally reversed and then determine whether a subsequently presented temporarily reversed version of the same word, or a closely related nonword, matched the temporally reversed representation. Compared to a rhyme judgment task, the word reversal task activated bilateral parietal regions similar to those known to be engaged in mental rotation of visual representations, rather than left inferior prefrontal regions known to be involved in phonological processing. This suggests that manipulation of phonological representations of lexical items in working memory is represented in the same neural regions as the manipulation of visual representations, indicating that similar cognitive processes may be involved. In tune with the results of Lehnert and Zimmer (2006), this study also suggests a common store for spatial representation of auditory and visual information, and thus provides further support for a working memory component such as the episodic buffer.

The episodic buffer and neuroimaging

More evidence of the nature of the episodic buffer comes from neuroimaging. Using fMRI, Bor and Owen (2007) studied neural regions supporting working memory processing requiring integration of information from perceptual sources with representations in long-term memory, which is one of the proposed functions of the buffer. They achieved this by manipulating the availability of strategic recoding opportunities during a working memory task. Strategic recoding was based either on inherent mathematical redundancy, i.e. the sequences to be remembered displayed numeric regularity, or well-established memories, i.e. the sequences to be remembered had been previously memorized by the participants. In both cases, behavioural evidence showed that availability of strategic coding made the task easier, suggesting that working memory processes were being supported by information held in long-term memory. It was found that activation of a prefrontal–parietal network was greater when strategic recoding could be applied than when it could not, irrespective of whether strategic recoding was based on mathematical redundancy or well-established memories. As strategic recoding facilitated performance, the prefrontal activation relating to strategic recoding could not be attributed to increased task difficulty, which is also known to activate the prefrontal cortex (Braver et al. 1997).

Other studies have addressed episodic buffer function in relation to binding of information that according to the original component model (Baddeley 1986; Baddeley and Hitch 1974) would be processed in the phonological loop and the visuospatial sketchpad. In an fMRI study (Prabhakaran et al. 2000) which investigated binding of phonological and spatial information in working memory, the bound condition involved maintaining letters and their location in a display and the separate condition involved maintaining letters at a neutral location along with a specific location indicated separately. Under both conditions, participants were instructed to memorize displayed information and recall was probed. The bound condition was less taxing than the separate condition and activated right frontal cortex more, while the separate condition preferentially activated posterior regions, including the temporal and occipital gyri bilaterally, and the cerebellum. These results were interpreted as demonstrating a frontal lobe component of episodic buffer processing. However, this interpretation assumes that the episodic buffer is concerned with *maintenance* of bound information rather than both *formation* and *maintenance* of bound representations, which is the functional definition given by Repovs and Baddeley (2006). Formation of bound representations takes place during the separate condition in the study by Prabhakaran et al. (2000) and thus the net activation for this condition in posterior regions appears to represent this function.

In a related fMRI study by Zhang et al. (2004), neural networks relating to the recall of auditory digits and visual locations in either mixed or separate order were investigated. It was hypothesized that recall of auditory digits and visual locations in mixed order required binding of order information across the auditory and visual modalities, and that this would tax the episodic buffer more than recall of separate order. It was shown that mixed order maintenance activated the right prefrontal cortex and the temporoparietal junction, irrespective of memory load. If mixed order recall is interpreted in terms of formation and maintence of unitary multidimensional representations in working memory, the finding agrees with the pattern of activation demonstrated by Prabhakaran et al. (2000). However, behavioural data showed that mixed order recall was also more taxing than separate order recall, and thus differential neural activation may also reflect executive processes. It is interesting to note that while results of the Lehnert and Zimmer (2006) study suggest a common store for *spatial* representation of auditory and visual information, the results of Zhang et al. (2004) suggest a common store for *order* representation of auditory and visual information.

In a recent study aimed at identifying the neural basis of executive function in working memory, Osaka et al. (2004) found evidence to suggest that the anterior cingulate cortex and the left inferior frontal gyrus were implicated. To date, there is no evidence to show that these particular regions are critically involved in episodic buffer processing. This further supports the case for a dissociation between the neural networks supporting the central executive and the episodic buffer.

None of these studies specifically addresses the role of the episodic buffer in language processing and further work needs to be done in this area, in particular, addressing the role of binding visual information in the service of language processing. In a recent study (Rudner et al. 2007a), however, we took a first step in this direction by studying the binding of lexical signs and words in working memory.

Imaging the episodic buffer for language processing

The episodic buffer plays a key role in language processing in binding multimodal input information with long-term memory representations (Rönnberg et al. 2007). In order to better understand this role, it is important to investigate neural underpinnings. In a recent study, we examined the neural representation of binding lexical items presented in the two languages of bilinguals with existing semantic representations in long-term memory (Rudner and Rönnberg 2006; Rudner et al. 2007b). We used a two back task in which stimuli were presented as audiovisual recordings of a model articulating lexical items in either one or other of the languages. The task was to determine for each item whether its meaning matched that of the item presented two steps back in the sequence in the other language. This meant that solution of the task demanded that the phonological surface form of each lexical item be bound to its underlying semantic representation stored in long-term memory. By contrast, in the two control conditions where stimuli were presented in only one language at a time, the task could be solved merely by storing and comparing surface representations. In order to polarize the surface forms of the two languages and thus enhance the requirement for binding phonological form to semantic representation, we chose languages in the two distinct language modalities of sign and speech, specifically Swedish Sign Language (SSL) and Swedish; the participants were native hearing signers. We found that binding phonological representations of lexical items to their underlying semantic representations activated a network of posterior regions including the right middle temporal lobe, and that activation of this network was associated with item-level processing. Activation of the right middle temporal lobe was interpreted as indicating binding of phonological surface representations of lexical items to their underlying semantic representations in working memory and that this was thus a signature of episodic buffer processing underpinning language.

The episodic buffer and other components of working memory

Another finding of this study was that episodic buffer processing seems to be supported by phonological loop processes; in this case associated with the sign loop. It has been suggested that the episodic buffer is associated with phonological loop processes (Repovs and Baddeley 2006). However, this is the first time, to our knowledge, that such an association has been shown. Three decades of research have demonstrated the importance of the phonological loop in relation to language processing and results are beginning to emerge regarding the role of the episodic buffer in language processing. Future work should focus on the interface of the phonological loop and the episodic buffer and examine how phenomena associated with the phonological loop; the phonological similarity effect, the word length effect and articulatory suppression interact with episodic buffer processes concerned with the binding of information in different codes from different perceptual and mnemonic sources.

Available data indicate that although episodic buffer processing may be associated with central executive processing, this is not necessarily the case. Imaging work suggests that separate networks may be involved in formation and maintenance of unitary, multidimensional representations in the episodic buffer with maintenance engaging frontal regions and formation engaging posterior regions. Activation of frontal regions in connection with episodic buffer processing has been interpreted as indicating a connection between the buffer and the central executive, which is also known to engage frontal regions (Repovs and Baddeley 2006). However, buffer processing that activates frontal regions has also been shown to be dissociated from general executive load (Bor and Owen 2007; Osaka et al. 2004; Prabhakaran et al. 2000). Evidence also suggests that buffer processes are engaged in effortful language processing under suboptimum conditions, and under such circumstances the episodic buffer and the central executive presumably work in tandem along with the phonological loop (Rönnberg et al. 2007). Future work should address the interface between the central executive and the episodic buffer and the way in which cognitive load interacts with formation and maintenance of unitary, multidimensional representations in the episodic buffer.

The role of the hippocampus in episodic buffer processing

In our study of the neural representation of the episodic buffer in a linguistic or communicative context (Rudner et al. 2007a), we found that the left hippocampus was engaged in binding phonological representations of lexical items to their underlying semantic representations at both item and task levels. A role for the hippocampus in episodic buffer processing is particularly interesting as this neural region which is known to play a key role in long-term memory encoding has, until recently, not been thought to play an important part in short-term memory processing (Cave and Squire 1992). However, that picture is beginning to change and a number of recent studies have shown that where binding is involved in working memory processing, the hippocampus is engaged. Davachi and Wagner (2002) showed that the hippocampus is more involved in the memorizing of lexical items when relational processing is involved. Mitchell et al. (2000) found that feature binding in working memory, which is also a function of the episodic buffer, activates the left hippocampus in young adults but not in older adults, who are known to have greater difficulty with feature binding (Chalfonte and Johnson 1996). Olson et al. (2006) found that hippocampal damage in neuropsychological patients was associated with an impairment of the ability to process object/location conjunctions in working memory but spared ability to process objects and locations separately. Petersson et al. (2006) analysed the functional and effective connectivity of interference induced by irrelevant sounds during a verbal working memory task and found a closer interaction between the verbal working memory system and the left hippocampus region during interference. They argued that interference from irrelevant sounds induces engagement of the episodic buffer and thus that the left hippocampus is involved in episodic buffer processing. This line of reasoning concurs with the argument that speech recognition in noise with input that is distorted in relation to long-term memory representations leads to engagement of the episodic buffer (Rönnberg et al. 2007). Moscovitch (1992) has proposed that a hippocampal component forms a link between working memory and long-term memory. This proposal predates introduction of the episodic buffer in the component model of working memory but is in line with recent findings that the hippocampus is engaged in integratory processes that have been functionally ascribed to the episodic buffer (Repovs and Baddeley 2006). Evidence indicates that both linguistic and non-linguistic processing in the episodic buffer leads to hippocampal engagement.

Findings of hippocampal engagement in episodic buffer processing are somewhat anomalous in view of the fact that, as we have seen, some amnesics, who presumably suffer from medial temporal lobe damage, have spared episodic buffer function (Baddeley and Wilson 2002; Gooding et al. 2005). This issue is illuminated by the work of Quinette et al. (2003, 2006) who studied working memory function during transient global amnesia (TGA). TGA is a neurological syndrome of uncertain etiology occurring in middle age and characterized by a profound, time-limited episodic memory impairment of acute onset. In an initial study (Quinette et al. 2003), it was found that while phonological loop, visuospatial sketchpad and central executive functions were spared during TGA, the pattern of results suggested that episodic buffer was compromised in some patients, who showed pathological immediate cued recall, but preserved in others, who scored normally on this task. In order to further investigate the role of the episodic buffer in TGA, Quinette et al. (2006) examined formation and maintenance of bound representations in working memory during the acute phase of TGA in a new group of patients. It was found that while formation and maintenance of multimodal representations were intact in patients with specific episodic memory *storage* disorder during the acute phase of TGA, these functions were impaired in patients with specific episodic memory *encoding* deficits. These results indicate that while episodic buffer processing is reliant on long-term memory encoding processes, it is dissociated from long-term memory storage.

Conclusions

We have reviewed behavioural, neuropsychological and neuroimaging data that support a component of working memory such as the episodic buffer postulated by Baddeley (2000) whose function is the formation and maintenance of unitary multidimensional representations. This function was originally attributed to the central executive of working memory, and early descriptions of the episodic buffer proposed that episodic buffer function would be closely related to central executive function both behaviourally and neurally. However, evidence indicates that the episodic buffer function is not necessarily dependent on central executive function and that the neural correlates of these two functions may be dissociated. We have discussed data indicating that the episodic buffer is involved in language processing and that its capacity may be important for language processing under suboptimum conditions where the language signal is poorly specified due either to sensory degradation or to a mismatch between working memory and long-term memory representations. We have also discussed data indicating that the formation and maintenance of unitary multidimensional representations on the basis of sign and speech is dependent on the right middle temporal lobe, probably reflecting semantic processes and that phonological loop processes seem to support these processes. Further we have reviewed neuropsychological and neuroimaging findings indicating a role for the hippocampus in episodic buffer processing and found an indication that it is specifically the encoding role of the hippocampus that is important in episodic buffer function.

These findings in favour of regarding the episodic buffer as a separate component of working memory within the framework of the component model (Baddeley 2000; Repovs and Baddeley 2006) can also be understood in terms of other models of working memory that emphasize the importance of the link between working memory and long-term memory. The ELU framework (Rönnberg 2003a; Rönnberg et al. 2007) introduced the concept of mismatch which arises when phonological representations in working memory do not readily match with long-term memory representations. Within the ELU framework, RAMBPHO serves to integrate multisensory, multilingual and long-term memory-based information, typically in an implicit stream. This function is equivalent to that of the episodic buffer in a communicative context. Continued investigation of episodic buffer function in relation to phonological, speed and capacity constraints will increase our understanding of working memory processes and their interface with long-term memory.

References

Allen RJ, Baddeley AD, Hitch GJ (2006) Is the binding of visual features in working memory resource-demanding? J Exp Psychol Gen 135:298–313

Alloway TP, Gathercole SE, Willis C, Adams AM (2004) A structural analysis of working memory and related cognitive skills in young children. J Exp Child Psychol 87:85–106

Andersson U, Lidestam B (2005) Bottom-up driven speechreading in a speechreading expert: the case of AA (JK023). Ear Hear 26(29):214–224

Baddeley A (1986) Working memory. Clarendon Press, Oxford

Baddeley A (1992) Working memory. Science 255:556–559

Baddeley A (1996) Exploring the central executive. Q J Exp Psychol A 49:5–28

Baddeley A (2000) The episodic buffer: a new component of working memory? Trends Cogn Sci 4(11):417–423

Baddeley A, Hitch G (1974) Working memory. In: Bower GA (ed) Recent advances in learning and motivation, Vol 8. Academic, New York, pp 47–90

Baddeley A, Wilson BA (2002) Prose recall and amnesia: implications for the structure of working memory. Neuropsychologia 40(10):1737–1743

Bayliss DM, Jarrold C, Baddeley AD, Gunn DM (2005) The relationship between short-term memory and working memory: Complex span made simple? Memory 13(3/4):414–421

Bor D, Owen AM (2007) A common prefrontal-parietal network for mnemonic and mathematical recodning strategies within working memory. Cereb Cortex 17(4):778–786

Boutla M, Supalla T, Newport EL, Bavelier D (2004) Short-term memory span: insights from sign language. Nat Neurosci 7(9):997–1002

Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC (1997) A parametric study of prefrontal cortex involvement in human working memory. Neuroimage 5:49–62

Buchsbaum B, Pickell B, Love T, Hatrak M, Bellugi U, Hickok G (2005) Neural substrates for verbal working memory in deaf signers: fMRI study and lesion case report. Brain Lang 95(2):265–272

Cave CB, Squire LR (1992) Intact verbal and nonverbal short-term memory following damage to the human hippocampus. Hippocampus 2:151–163

Chalfonte BL, Johnson MK (1996) Feature memory and binding in young and older adults. Mem Cogn 24(4):403–416

Daneman M, Carpenter PA (1980) Individual differences in working memory and reading. J Verbal Learn Verbal Behav 19:450–466

Davachi L, Wagner AD (2002) Hippocampal contributions to episodic encoding: insights from relational and item-based learning. J Neurophysiol 88:982–990

Ericsson KA, Kintsch W (1995) Long-term working memory. Psychol Rev 102:211–245

Foo C, Rudner M, Rönnberg J, Lunner T (2007) Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. J Am Acad Audiol 18:553–566

Ganis G, Thompson WL, Kosslyn SM (2004) Brain areas underlying visual mental imagery and visual perception: an fMRI study. Cogn Brain Res 20(2):226–241

Gathercole SE, Pickering SJ, Knight C, Stegmann Z (2004) Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. Appl Cogn Psychol 18(1):1–16

Gooding PA, Isaac CL, Mayes AR (2005) Prose recall and amnesia: more implications for the episodic buffer. Neuropsychologia 43:583–587

Jeffries E, Lambon Ralph MA, Baddeley AD (2004) Automatic and controlled processing in sentence recall: the role of long-term and working memory. J Memory Lang 51:623–643

Just MA, Carpenter PA (1992) A capacity theory of comprehension: individual differences in working memory. Psychol Rev 99(1):122–149

Klauer KC, Zhao Z (2004) Double dissociations in visual and spatial short-term memory. J Exp Psychol Gen 133(3):355–381

Lehnert G, Zimmer HD (2006) Auditory and visual spatial working memory. Mem Cogn 34(5):1080–1090

Logie RH, Della Sala S, Wynn V, Baddeley AD (2000) Visual similarity effects in immediate verbal serial recall. Q J Exp Psychol A 53(3):626–646

Lunner T (2003) Cognitive function in relation to hearing aid use. Int J Audiol 42(Suppl 1):S49–S58

Lyxell B, Arlinger S, Andersson J, Bredberg G, Harder H (1998) Phonological representation and speech understanding with cochlear implants in deafened adults. Scand J Psychol 39:175–179

Mitchell KJ, Johnson MK, Raye CL, D'Esposito M (2000) fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. Cogn Brain Res 10:197–206

Moscovitch M (1992) Memory and working-with-memory: A component process model based on modules and central systems. J Cogn Neurosci 4(3):257–267

Olson IR, Page K, Moore KS, Chatterjee A, Verfaellie M (2006) Working memory for conjunctions relies on the medial temporal lobe. J Neurosci 26(17):4596–4601

Osaka N, Osaka M, Kondo H, Morishita M, Fukuyama H, Shibasaki H (2004) The neural basis of executuive function in working memory: an fMRI study based on individual differences. Neuroimage 21:623–631

Petersson KM, Gisselgård J, Gretzer M, Ingvar M (2006) Interaction between a verbal working memory network and the medial temporal lobe. Neuroimage 33:1207–1217

Prabhakaran V, Narayanan K, Zhao Z, Gabrielli JDE (2000) Integration of diverse information in working memory in the frontal lobe. Nat Neurosci 3:85–90

Quinette P, Guillery B, Desgranges B, de la Sayette V, Viader F, Eustache F (2003) Working memory and executive functions in transient global amnesia. Brain 126:1917–1934

Quinette P, Guillery B, Noël A, de la Sayette V, Viader F, Desgranges B, Eustache F (2006) The relationship between working memory and episodic memory disorders in transient global amnesia. Neuropsychologia 44:2508–2519

Ranganath C, Blumenfeld RS (2005) Doubts about double dissociations between short- and long-term memory. Trends Cogn Sci 9(8):374–380

Repovs G, Baddeley A (2006) The multi-component model of working memory: explorations in experimental cognitive psychology. Neuroscience 139(1):5–21

Rönnberg J (1993) Cognitive characteristics of skilled tactiling: the case of GS. Eur J Cogn Psychol 5:19–33

Rönnberg J (2003a) Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. Int J Audiol 42:S68–S76

Rönnberg J (2003b) Working memory, neuroscience and language: evidence from deaf and hard of hearing individuals. In: Marschark M, Spencer P (eds) The handbook of deaf studies, language, and education. Oxford University Press, Oxford

Rönnberg J, Rudner M, Ingvar M (2004) Neural correlates of working memory for sign language. Cogn Brain Res 20:165–182

Rönnberg J, Rudner M, Foo C (2007) The cognitive neuroscience of signed language: applications to a working memory system for sign and speech. In: Bäckman L, Nyberg L (eds) Memory, aging and the brain: a Festschrift in honour of Lars-Göran Nilsson. Psychology Press, London

Rossi-Arnaud C, Pieroni L, Baddeley A (2006) Symmetry and binding in visuo-spatial working memory. Neuroscience 139:393–400

Rudner M, Rönnberg J (2006) Towards a functional ontology for working memory for sign and speech. Cogn Process 7(5):183–186

Rudner M, Rönnberg J, Hugdahl K (2005) Reversing spoken items: mind twisting not tongue twisting. Brain Lang 92(1):78–90

Rudner M, Fransson P, Ingvar M, Nyberg L, Rönnberg J (2007a) Neural representation of binding lexical signs and words in the episodic buffer of working memory. Neuropsychologia 45(10):2258–2276

Rudner M, Foo C, Rönnberg J, Lunner T (2007b) Phonological mismatch makes aided speech recognition in noise cognitively taxing. Ear Hear (in press)

Tulving E (1983) Elements of episodic memory. Oxford University Press, Oxford

Vallar G, Baddeley AD (1987) Phonological short-term store and sentence processing. Cogn Neuropsychol 4:417–438

Vass M, Ibertsson T, Lyxell B, Sahlén B, Hällgren M, Larsby B (2006) Cognitive development in children with cochlear implants (CI) ASHA Convention, Miami Beach, November 16–18

Wilson M, Emmorey K (1997) A visuospatial "phonological loop" in working memory: evidence from American sign language. Memory Cogn 25:313–320

Wilson M, Emmorey K (1998) A "word length effect" for sign language: further evidence for the role of language in structuring working memory. Memory Cogn 26(3):584–590

Wilson M, Emmorey K (2003) The effect of irrelevant visual input on working memory for sign language. J Deaf Stud Deaf Educ 8:97–103

Zhang D, Zhang X, Sun X, Li Z, Wang Z, He S, Hu X (2004) Cross-modal temporal order memory for auditory digits and visual locations: an fMRI study. Hum Brain Mapp 22(4):280–289