RUNNING HEAD: Working Memory and Language

Working memory and (second) language processing

Arnaud Szmalec, Marc Brysbaert, Wouter Duyck

Ghent University, Belgium

Contact:

Arnaud Szmalec

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

B-9000 Ghent (Belgium)

arnaud.szmalec@UGent.be

Abstract

This chapter discusses the interactions between two of the most important human cognitive functions: memory and language. First, the concept of working memory is introduced, along with a brief summary of the evolutions that working memory theory has undergone in the last decades. The second part of the chapter focuses on the role of (verbal) working memory in language acquisition and processing. It is argued that working memory, and especially the ability to temporarily represent serial-order information, is crucially involved in both native and foreign word learning, and perhaps also in sentence and text comprehension. The third and final part of the chapter explores the other direction of the interaction, by questioning whether language processing can influence working memory functioning. This question is addressed with recent behavioral and neurological evidence for a general executive control advantage in bilinguals, which makes a strong case for the trainability of some aspects of working memory.

Keywords: working memory, language acquisition, language processing, executive control, bilingualism

Working memory and (second) language processing

From intelligence testing to working memory

In 1887, Jacobs published a series of studies in which he reported that older children could repeat longer strings of digits read out to them than younger children. Jacobs also reported that intelligent children (as assessed by the teacher) could repeat more digits than less intelligent children. This idea was picked up by Binet and Simon in the early 20th century when they developed the first valid intelligence test. They found that 3-year old children could repeat only sequences of two digits, whereas children of 4 years could repeat sequences of three digits, and most healthy children of 7 years could repeat sequences of five digits. Therefore, Binet and Simon included digit repetition in their intelligence test (Binet & Simon, 1905). Ever since, the digit span task (as it became called) has been part of intelligence tests, because it correlates reasonably well with the scores of other subtests of intelligence (such as arithmetic, general information, and the discovery of similarities). The task received further impetus when Miller (1956) argued it was a good measure of a person's short-term memory capacity.

In the early 1970s several authors felt uneasy with the digit-span as a measure of memory capacity. It seemed to consider short-term memory too much as a passive storage buffer, rather than an active part of human information processing. As a result, the concept of working memory, representing both storage and executively controlled manipulation of information, was put forward. An important publication in this respect was the working memory model of Baddeley and Hitch (1974). This model consisted of three parts: (i) a modality-free central executive related to attention, (ii) a phonological loop holding information in a speech-based form, and (iii) a visuo-spatial sketchpad for the coding of visual and spatial information. A further milestone was the publication by Daneman and

Carpenter (1980) of an article in which they presented the reading span task as a measure of working memory capacity. This task (also known as the complex span task) was developed to simultaneously tax the storage and processing functions of working memory. Participants had to read sentences (the processing component) while maintaining and retrieving the final words of the sentences (the storage component). An example of a test item with two sentences was:

- When at last his eyes opened, there was no gleam of triumph, no shade of anger.

- The taxi turned up Michigan Avenue where they had a clear view of the lake.

After reading aloud these two sentences the participant had to retrieve the two last words (anger, lake). The number of sentences was increased until the participants made errors. Daneman and Carpenter (1980) observed that reading span typically varied from 2 to 5 words. They further discovered that this span correlated much better with reading comprehension and performance on the Scholastic Aptitude Test (SAT; a standardized test for college admission in the US) than the traditional, passive word span (measured by presenting lists of words of varying length to participants and asking them to repeat the lists). Subsequent reviews confirmed the high correlations between working memory capacity and language comprehension (Daneman & Merikle, 1996), and between working memory capacity and fluid intelligence (Ackerman, Beier, & Boyle, 2005; Engle, Tuholski, Laughlin, & Conway, 1999).

A wide variety of immediate serial recall tasks and complex span measures are in use today to increase our understanding of the structure and the functioning of working memory. At the same time, the conceptualization of working memory has gone through some substantial changes. Whereas Daneman and Carpenter (1980) considered working memory as a unitary system with a single capacity, later research provided evidence for several subcomponents with their own capacities. For instance, Jarrold and Towse (2006), in line with

Baddeley and Hitch's (1974) model, argued that working memory capacity depended on (i) processing efficiency, (ii) storage capacities for the maintenance of verbal/numerical information and spatial information, and (iii) controlled attention needed for the coordination and integration of storage and processing, and for the inhibition of irrelevant information.   Still other working memory theorists started to question the idea of working memory as a separate module. Partially inspired by the work of Cowan (1988), which was further elaborated by Oberauer (2009), they have questioned the multiple-component view of memory and argued that short-term memory, long-term memory, and working memory are not separate structures but differ from each other in terms of activation levels of representations in memory and the amount of attentional control dedicated to those representations. Working memory then is seen as an activated subset of long-term memory, with information in a directly accessible state and shielded against interference from other memory contents through attentional control (e.g., Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011). In this view, the same memory processes (recall, recognition, recollection, stimulus familiarity,etc..) operate in the entire memory system (e.g., Goethe, & Oberauer, 2008; Oztekin & McElree, 2007) and the structural differentiation between memory subsystems is largely abandoned. This assumption is based on a wide variety of behavioral and neurophysiological evidence showing that long-term and working memory are in much closer interaction than initially thought (which also led to the introduction of the Episodic Buffer in the traditional working memory model; Baddeley, 2000). Probably one of the best examples of the close collaboration between working memory and long-term memory is language acquisition.

How working memory supports language acquisition

Why a Working Memory for Verbal Information?

A longstanding tradition in memory research draws a distinction between verbal and visuospatial information, on which different memory processes operate. This division was explicitly present in the working memory model of Baddeley and Hitch (1974) and it remains present in many recent models. The need for a visuospatial (working) memory is easy to assert on the basis of evolutionary grounds. A visuospatial working memory provides a survival advantage because it allows for the retention of visual and spatial information when the information is no longer accessible from the sensory registers (see Vandierendonck and Szmalec, 2011, for a collection of recent papers on spatial working memory theory). Such a memory system allows an organism, for example, to keep track of its nest and to remember where prey is hiding or where predators are likely to be waiting.

Scientists have sought for similar evolutionary factors behind the verbal part of (working) memory: Why did humans evolve the capacity to (briefly) retain speech-based information? What is the evolutionary purpose of working memory for speech-based materials? It seems unlikely that verbal working memory developed to perform well in cognitive psychology experiments, where participants are asked to recall lists of digits, syllables, telephone numbers, words or other artificial stimuli cognitive psychologists are interested in. What could be the etiology of verbal working memory?

An important breakthrough was published in a seminal paper by Baddeley, Gathercole, and Papagno (1998). On the basis of a literature review, they proposed that verbal working memory primarily represents "the processes and mechanisms by which the sound patterns of the words of the native language are learned by the child" (p. 159). Similar perspectives on verbal working memory had been introduced before, as in the work of Martin and Saffran (1992) who suggested that short-term memory for verbal information was merely an emergent property of the temporary activation of linguistic information. However, the conceptualization

of Baddeley et al. (1998) was more dynamic and open to individual differences, as typically investigated in working memory research. As a result, Baddeley et al. (1998) have been highly influential and stimulated a lot of new research, which is summarized below.

Verbal Working Memory and the Learning of New Words

Baddeley et al. (1998) reviewed a large amount of evidence from adults, children, and patients in support of the idea that verbal working memory primarily is a language learning device. For example, positive correlations were reported between measures of verbal working memory capacity (e.g., nonword repetition) and native vocabulary knowledge in children of various ages (Bowey, 2001; Gathercole & Adams, 1993, 1994). Further experimental evidence was found by Gathercole and Baddeley (1990) when they taught 5- to 6-year-old children to learn unfamiliar names (e.g., Pimas) to new toy animals. This was considered as an operationalization of naturalistic *word* learning, namely the mapping of a new word form to a referent in the real world. As expected, performance was worse for children with low nonword repetition scores than for children with high scores. Further experiments tried to make sure that the correlation between memory performance and learning new words was really due to constraints imposed by working memory capacity on language learning, and not due to a third, confounded variable, or to the fact that the working memory capacity of an individual depends on the language learning skills of the individual (see also Gathercole, 2006). A first line of research (Duyck, Szmalec, Kemps, & Vandierendonck, 2003; Papagno, Valentine, & Baddeley, 1991) used the dual-task methodology to demonstrate that loading verbal working memory resulted in poorer learning of word-nonword pairs, such as finger-*vilsan* (in which participants had to name the nonword upon hearing the word, or vice versa),

but not in learning word-word pairs (e.g., frog-nail), in both adults and 11- to 13-year-old children.

Other research examined patients with verbal short-term memory deficits, to see whether these patients found it hard to learn new vocabulary, while at the same time retaining the language capacities they had before the lesion. For instance, Baddeley (1993) reported a case study of such a patient, SR, who indeed performed very poorly on a word-new word association task (pairing English with Finnish words), unless he could associate both words by forming very elaborate semantic associations. There is also good evidence that children with specific language impairment have a reduced verbal working memory capacity and that the latter is causally linked to language acquisition difficulties (Baddeley et al., 1998).

The Importance of Serial Order Memory for Novel Word Learning

More recent research has focused on the question *what* must be learned in a word-new word association task (different from a word-word association task) and how working memory is related to this. Is verbal working memory especially important for learning the new word form itself, for associating the old with the new word form, or for mapping the semantic representation of the old word form to the new word form? Or could it be that the effect of verbal working memory on word – new word association is simply due to the fact that a new, conflicting name must be given to existing information, which already has a name? In the latter case, working memory would play a role in the learning of a new language (where existing objects get new names) but not in the learning of new names of new objects.

An important model to describe the acquisition of new names for new objects (as happens in children acquiring language) was proposed by Page and Norris (1998). The rationale behind their Primacy Model of immediate serial recall can be summarized as

follows. Learning a novel word form consists of learning both a sequence of sounds (letters) and the correct order of the sounds. Learning the new word *artecey*, for example, involves learning both the identity and the order of the phonemes: *ar, te, cey*. According to Page and Norris (2009), learning such a word is similar to learning the sequence of letters in a letter span task (i.e., repeating the letters R T C in an immediate serial recall task). In other words, the model of Page and Norris bridges the gap between verbal working memory and language acquisition by hypothesizing that the working memory mechanisms involved in immediate serial recall of letters are the same as those involved in the acquisition of novel word-forms. In this view, naturalistic word-form learning consists of extracting regularities from the auditory information in the environment. If a baby repeatedly hears the sequence *ar*, *te, cey* in this specific order, it will develop a lexical representation for *artecey*, which may then be linked to a real-world referent that is always present when the baby hears this particular sequence of sounds.

Page and Norris (2009) argued that the above naturalistic word-learning process can be mimicked in a laboratory setting using the Hebb repetition effect (Hebb, 1961). The Hebb repetition effect is observed in an immediate verbal serial recall task when a particular sequence of digits/syllables is repeated across trials. In such a situation, recall of the repeated sequence improves over time relative to that of unrepeated sequences. In essence, Page and Norris (2009) argued that the Hebb repetition effect showed how information related to a sequence of items (like letters or syllables) in working memory gradually develops into a stable long-term memory trace that has the same characteristics as a newly acquired word form. The first empirical evidence supporting Page and Norris's computational exercise was reported by Mosse and Jarrold (2008). They found a positive correlation between the steepness of the Hebb learning curve and performance in a paired-associate learning task with

nonwords, in a sample of 5- to 6-year olds. More recently, Szmalec, Duyck, Vandierendonck, Barbera Mata and Page (2009) presented the first experimental demonstration of the idea that the verbal Hebb effect can be used as a laboratory analogue of novel word-form learning. They presented adult participants with sequences of syllables in a standard Hebb learning paradigm (e.g., *zi-lo-ka-ho-fi-se-be-ru-mo*). Then, the same participants took part in a lexical decision experiment including nonwords that were constructed with syllables from the Hebb experiment (*ziloka, hofise, berumo*). Szmalec and colleagues observed that participants were slower to reject the Hebb-based nonwords, compared to matched control nonwords. This suggests that the immediate serial recall of repeated Hebb sequences leads to representations in lexical memory similar to those of existing words, just as would happen when people acquire novel words.

The idea that the Hebb repetition effect operationalizes the memory mechanisms that support language learning raises the question whether the Hebb effect can also be used to shed some light on language impairment. In this context, Szmalec, Loncke, Page and Duyck (2011) demonstrated that adults with dyslexia show impaired Hebb learning across verbal and visuospatial stimulus modalities. On the basis of these findings, they put forward a new, memory-based account of dyslexia, in which the various difficulties experienced by people with dyslexia are assumed to originate from an impairment affecting the learning of serial-order information in memory, of which Hebb repetition learning is an example. Assuming that a newly learned word-form is simply an ordered sequence of sublexical items, the Hebb learning account of dyslexia proposes that the lexical representation's constituent elements are not optimally consolidated as a single entry in long-term memory. Hence, lexical access for that entry during reading will be impaired and normal procedures for mapping grapheme sequences to phoneme sequences are disrupted (Whitney & Cornelissen, 2005).

Similar ideas have been proposed by Gupta (2009), who also sees a new word (or a nonword) as a novel sequence of sounds, similar to a sequence of digits or letters to be learned in an immediate serial recall task (or span task). Based on this assumption, Gupta hypothesized that effects typically observed in immediate serial recall tasks should be present in the learning of new words or nonwords as well. In a series of elegant studies, he indeed observed that the phonemes of syllables within a newly learned word form are subject to primacy and recency effects (Gupta, 2005; Gupta, Lipinski, Abbs, & Lin, 2005). The primacy effect refers to the finding that the items presented first in a series are better recalled than items presented later, independent of whether the test follows immediately after the presentation of the series or after a distraction task that depletes short-term memory. The recency effect refers to the finding that the items presented last are better recalled when the test immediately follows the series presentation but not when a distraction task intervenes. The fact that the same effects are observed in new word learning and serial recall further strengthens the claim that naturalistic word-form learning and immediate serial recall rely on the same working memory mechanisms. Based on these findings, Gupta (2009; Gupta & Tisdale, 2009) developed a computational model that could simulate the various effects by making a distinction between a lexical (word) level and a sublexical (sound) level of item representations and a serial order mechanism that encoded the order of the lexical and sublexical elements.

A third model stressing the analogy between serial-order learning in short-term memory tasks and language learning was developed by Burgess and Hitch (1999, 2006). An interesting feature of this model is that it is a connectionist model consisting of nothing but layers of nodes connected to each other. This allowed the authors to provide a common explanatory mechanism for effects like serial position, lexicality and Hebb repetition. Further

important is that the model made an explicit distinction between serial order information for unknown stimuli (new words) and item information for known stimuli (old words), which were based on different processes.

Finally, the distinction between memory for item information and memory for order information has been documented in neuroscientific and neuropsychological research as well (Majerus, Lekeu, Van der Linden, & Salmon, 2001; Majerus, Poncelet, Elsen, & Van der Linden, 2006; Majerus, Poncelet, Greffe, & Van der Linden, 2006; Majerus, Van der Linden, Mulder, Meulemans, & Peters, 2004). Using a correlational approach, Majerus et al. (2006) explored the contribution of three different short-term memory skills to novel word-form learning in patients. These were short-term memory for serial order information, item recall, and item recognition. The results showed that only memory for serial order played a role in acquiring novel phonological word forms and, therefore, supported the hypothesis that the representation of item and order information are distinct factors in word learning. Majerus et al. (2006) further explored the item v. order distinction with fMRI. They observed that memory for order and items activated different brain regions. Order relied on the right intraparietal sulcus, the right cerebellum, and the bilateral premotor cortex, whereas item memory activated two regions associated with language processing, namely the superior temporal gyrus and the left fusiform gyrus.

Overall the above findings provide compelling evidence for a causal relation between short-term serial recall and naturalistic word-form learning. Therefore, they reinforce the assumption that the primary purpose of human verbal working is to support the acquisition of language.

Verbal Working Memory and Second Language Word Learning

Evidently, the findings described in the previous sections have implications for second language (L2) learning. One of the key requirements of L2 learning is the acquisition of new word forms, which initially are nothing but sequences of sounds and letters.

Service (1992) was one of the first to specifically examine the relationship between nonword repetition and learning new words in L2. She ran a longitudinal study of Finnish-speaking primary school children learning English. At the beginning of the study, a nonword repetition task was administered and the scores on this test were correlated with English performance levels nearly three years later. Service observed that the nonword spans were a significant, independent predictor of L2 proficiency. Cheung (1996) ran another early study. He correlated nonword span with the number of trials 7th grade participants from Hong Kong needed to acquire new English L2 words. Cheung found the expected inverse relationship (participants with higher nonword spans learned the words faster), at least for the participants with vocabulary sizes lower than average, in line with the idea that verbal working memory is particularly important for acquiring new words and less so for the processing of familiar words.

The studies of Service (1992) and Cheung (1996) have since been replicated and extended in several studies with convergent results, a good review of which is given by Hummel and French (2010). So, there is little doubt that verbal working memory is involved in the acquisition of L2 words as much as it is in the acquisition of new L1 words. It also seems reasonable to assume that the working memory processes involved in L2 and L1 word learning are the same, although there is not much empirical evidence on this aspect yet, except for a study by Majerus, Poncelet, Van der Linden, and Weekes (2008). At the same time, there is fMRI evidence that for low-proficiency bilinguals order encoding may be less

efficient in L2 than in L1 (Majerus, et al, 2008), suggesting that in early stages L2 word

learning may be more difficult than L1 word learning.

Working Memory Involvement in Other Aspects of Language Processing

So far, we have reviewed evidence showing that verbal working memory (more precisely

memory for serial order and item information) supports the acquisition of novel lexical forms,

both in native and foreign languages. It is important to realize, however, that hypotheses about

the involvement of working memory in the human language system have not been restricted

to word learning. In the final section of this part, we briefly review some more ideas that have

been proposed about how working memory may be involved in the integration of individual

words into coherent sentences and discourse representations. Indeed the correlation between

working memory span measures and reading comprehension, originally discovered by

Daneman and Carpenter (1980), strongly points to the importance of working memory for text

understanding. However, it has been very difficult thus far to design paradigms that are as

convincing as those of novel word acquisition.

One of the first questions addressed by working memory proponents was whether

working memory is involved in sentence parsing (Just & Carpenter, 1992; Waters & Caplan,

1996). Sentence parsing refers to the processes needed to organize the words of a sentence

into a proposition (or set of propositions) summarizing who did what to whom. Indeed, it

seems obvious that verbal working memory (or the phonological loop in Baddeley and

Hitch's model) is needed to retain the surface structure of a sentence until the proper syntactic

interpretation has been made. Sentences can be syntactically complex with large distances

between related parts (e.g., between the subject and the verb, as in "when the girl with the red

hood, who was dancing in the wood, saw…"). In addition, many sentences are locally

ambiguous and may require some kind of reanalysis. This is shown most clearly in so-called

garden-path sentences, such as "the horse chased past the barn fell". For these sentences,

participants are likely to experience parsing difficulties because the structure of the sentence

does not agree with the initially preferred interpretation (i.e., "the horse that was chased" vs.

"the horse that was chasing"). Given the need to retain word order information until the

correct syntactic interpretation has been found, it seems reasonable to assume that people with

high working memory capacity will perform better on sentence parsing than people with low

capacity (e.g., Swets, Desmet, Hambrick, & Ferreira, 2007; Vallar & Baddeley, 1984).

A problem with this intuitively appealing hypothesis, however, is that syntactic

comprehension seems to be affected little by neurological conditions resulting in reduced

working memory capacity. Only for very complex sentences can an effect be shown. This

finding led Caplan and Waters (1999) to argue that sentences are interpreted by a system

independent of working memory (the so-called separate sentence interpretation resource),

giving rise to a vivid discussion about whether or not verbal working memory as traditionally

measured is needed for sentence parsing (e.g., Lauro, Reis, Cohen, Cechetto, & Papagno,

2010, for a recent installment). O'Brien, Segalowitz, Collentine, & Freed (2006), for instance,

claimed that verbal working memory capacity (as measured with nonword repetition) predicts

the development of narrative and grammatical competences in L2, in English-speaking adults

learning Spanish. Still, it cannot be denied that the consequences of reduced memory span are

much more severe for novel word learning than for sentence parsing, suggesting that the

involvement of verbal working memory will be less for the understanding of sentences than

for the learning of new words.

In a review article on the relationship between working memory and language,

Baddeley (2003) mentioned two other possible contributions of working memory to language

understanding. He thought it likely that visuospatial working memory would be involved in

maintaining a representation of the page and its layout during reading. Readers are amazingly

accurate at localizing previously read words. This can be seen, for instance, when they make

regressions upon encountering a comprehension problem in text reading. These regressive eye

movements are usually remarkably accurate (Kennedy, Brooks, Flynn, & Prophet, 2003) and

seem to require access to a spatial map of the text. Baddeley (2003) further hypothesized that

visuospatial working memory may also be involved in the understanding of spatial

information (e.g., grammatical structures involving spatial terms such as above, below,

shorter, and so on). Both ideas, however, still need to be tested.

How language processing supports working memory

Thus far, we have summarized findings showing that working memory is crucially involved

in language acquisition. Of equal interest is the reverse question, namely whether working

memory's processing (executive control) and storage (span/capacity) functions are also

influenced by language processing, or whether they remain unchanged. One research area that

has proven particularly fruitful in this respect concerns the consequences of bilingualism for

executive control functions.

Executive Control Advantages in Bilingualism

Recent studies point towards important cognitive benefits of being bilingual. Bialystok, Craik,

and Freedman (2007), for example, found that the age of onset of dementia is on average four

years later in bilinguals than in monolinguals. Cognitive advantages of bilingualism are

assumed to originate from the requirement to continuously control the activation of lexical

representations from the non-target language so that they do not interfere with the ongoing

language processing (Green, 1998).

There is now a good consensus that both languages of a bilingual are always to some extent active in lexical memory and interact with each other (Brysbaert & Duyck, 2010). For instance, it has been observed that bilinguals read native language (L1) words faster if the L2 translations are similar in form (i.e., when they are so-called cognates, such as apple and *appel* in English and Dutch). This is even true when the participants are reading complete sentences in L1 (Van Assche, Duyck, Hartsuiker, & Diependaele, 2009), which indicates that lexical access in bilinguals is not language selective. Similarly, there is evidence for unremitting competition between word forms from different languages in speech production. Ivanova and Costa (2008), for example, reported that L1 speech production is slower in Spanish-Catalan bilinguals than in monolinguals. Gollan and Acenas (2004) observed that bilinguals experience more tip-of-the-tongue states than monolinguals (these are situations in which one cannot retrieve the correct lexical entry for a concept).

Despite the fact that the languages of a bilingual are constantly in competition with each other, there is little evidence for control *failures*, as can be concluded from the few switching errors made. These are rare in comparison with other types of errors and hesitations in speech. Hence, bilinguals seem to have an efficient cognitive control mechanism dealing with the language competition in a highly interactive bilingual language processing system. This raises questions about the nature and the functioning of such a cognitive control system, and the extent to which this control system is specialized for language, or generalizes to other cognitive domains.

Interest in the language control of bilinguals took off after the publication of papers by Meuter and Allport (1999) and Costa and Santesteban (2004). In the former study, bilinguals were required to name pictures in the language indicated by an external cue. The experiment contained trials in which the language was the same as in the previous trial, and trials in

which the language switched. Meuter and Allport observed that bilinguals were slower in the switch trials than in the non-switch trials, and that the switching costs were larger when the language changed from L2 to L1 than vice versa. They interpreted the latter finding as due to the fact that more inhibition of L1 is required when participants speak in L2 than the other way around.

Costa and Santesteban (2004) adopted the same paradigm, and additionally manipulated second-language proficiency in a group of Spanish-Catalan bilinguals. For unbalanced bilinguals, they replicated the findings of Meuter and Allport (1999): a cost when switching languages, and a larger switching cost when switching to the dominant language. However, for balanced Spanish-Catalan bilinguals (who used both languages interchangeably and equally often), the language-switching cost was symmetrical, equally large in both directions. Surprisingly, the switching cost remained symmetrical when in a later experiment switching between L1 and a much-weaker L3 was studied. On the basis of this finding Costa and Santesteban argued that balanced bilinguals develop a qualitatively different mechanism of lexical selection, which can also be used for a weaker language (L3). Similarly, Costa, Santesteban and Ivanova (2006) found no asymmetrical switching cost in balanced Spanish-Catalan bilinguals who were switching between L2 and L3. These participants did show an asymmetry, however, when they were asked to switch between L3 and L4, leading Costa and colleagues to argue that there are limits to the extent to which the specific control mechanism can be applied.

All in all, it looks like the continuous language control and the repeated practice of language switching provide bilinguals with efficient control mechanisms, which generalize to some extent beyond the specific language pair for which they are needed (see the symmetrical language switching cost for L3 in Costa & Santesteban, 2004). This has led researchers to

investigate whether bilinguals also perform differently, more efficiently, in tasks that do not require verbal processing.

Within this framework, Bialystok and colleagues compared bilinguals with monolinguals on a wide variety of tasks tapping into executive control, such as the Simon task. In this task participants are required to make a spatial response (e.g., press the left or right key) to a non-spatial characteristic (e.g., the green or red color) of a stimulus, which has a particular spatial position (e.g., left or right of the fixation location). A robust finding in this paradigm is that the irrelevant position of the stimulus interferes with the response, such that participants are faster to respond to congruent trials (pressing the left key to a stimulus presented left of fixation) than to incongruent trials (pressing the right key to a stimulus presented left of fixation). This is the so-called Simon-effect. Interestingly, Bialystok, Craik, Klein and Viswanathan (2004) found a smaller Simon-effect in English-Tamil bilinguals than in monolinguals. Similarly, Bialystok, Craik and Ryan (2006) observed that bilinguals had fewer problems to move their eyes in the direction opposite to the one where the stimulus appeared (e.g., to move the eyes to the right when a light flash appears to the left; this is the so-called antisaccade task). Emmorey, Luk, Pyers and Bialystok (2008) found a comparable bilingual executive control advantage using a flanker task, in which irrelevant flanking stimuli had to be ignored for good task performance. The finding that bilinguals show better performance than monolinguals on a variety of executive control tasks suggest that the cognitive benefits of bilingualism are not restricted to language control. They generalize to other situations in which a dominant response must be suppressed for good performance.

The cognitive mechanisms underlying the bilingual executive control advantage were further explored by Prior and Gollan (in press), who manipulated the degree of language switching in their participants. They compared a group of balanced Spanish-English

bilinguals who regularly switched between languages, to a group of balanced Mandarin-English bilinguals who switched between languages less often. Only the Spanish-English bilinguals showed a reduced task switching cost. On the basis of this finding, Prior and Gollan concluded that only bilinguals who often switch between languages train their executive control capacities, causing improved task switching performance.

Also at the neural level there is evidence that language control makes use of the same brain structures as other, non-verbal control mechanism. Pioneering work was done by Hernandez and colleagues (Hernandez, Martinez, & Kohnert, 2000; Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001), who used a picture naming task. They observed stronger activation in the dorsolateral prefrontal cortex when bilinguals had to switch response languages. Using event-related fMRI, Wang, Xue, Chen, Xue and Dong (2007) reported stronger activation in the frontal cortices bilaterally and in the left anterior cingulate cortex when Chinese-English bilinguals switched picture naming from L1 to L2, but not from L2 to L1. The dorsolateral prefrontal cortex also showed stronger activation on switch trials, but this pattern of activation did not interact with switching direction. Similarly, a Positron Emission Tomography study by Crinion, et al (2006) identified involvement of the basal ganglia (caudate) in language control in German-English and Japanese-English bilinguals. Finally, Abutalebi and Costa (2008) reported that that naming in the first language in a bilingual context (compared with monolingual contexts) increased activation in the left caudate and anterior cingulate cortex. The brain areas that have been found to be involved in such language switching tasks highly overlap with the neural circuits identified in domain-general executive control research (Brass & von Cramon, 2002, 2004; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). Recent models of the anterior cingulate cortex assume that it is involved in conflict processing (Botvinick, Braver, Barch, Carter, & Cohen,

2001). Also, the basal ganglia are known to be crucial for cognitive flexibility (Aron, Watkins, Sahakian, Monsell, Barker, & Robbins, 2003).

Working Memory Capacity and Bilingualism

Another question that has been raised is whether bilingualism can boost total working memory capacity. Is it the case that bilinguals have a higher working memory capacity than monolinguals?

Some of the early research indicated that multilingual adults indeed have greater short-term memory and working memory spans (as measured with digit and nonword repetition) than IQ-matched monolingual participants. Papagno and Vallar (1995) compared Italian polyglots and controls on a series of tasks, related to fluid intelligence (Raven Progressive Matrices), L1 vocabulary (the WAIS subtest), auditory digit span, nonword repetition, visuospatial span (Corsi blocks), visuospatial learning (the number of sequences participants needed to learn a supra-span sequence on the Corsi blocks), paired-associate learning of words, and paired-associate learning of words and non-words. Multilinguals had similar or inferior performance to the controls for fluid intelligence, L1 vocabulary, visuospatial span, visuospatial learning, and paired-associate word-word learning. In contrast, they had better performance on digit and nonword repetition and on paired-associate word-nonword learning. Furthermore, a principal component analysis indicated that the latter three tasks were part of the same component (i.e., the participants' scores on these tests correlated substantially).

Not all findings have been positive, however. French and O'Brien (2008) tested Arabic and English non-word repetition before and after French-speaking children took part in an intensive English-as-a-second-language program. Whereas performance on the English nonwords improved (as expected from the finding that the memory span is larger for

meaningful words than for nonwords), no difference was observed for Arabic nonword repetition. Similarly, Vejnovic, Milin, and Zdravkovic (2010) compared reading spans of Serbian-English bilinguals in L1 and L2 as a function of L2 proficiency. Whereas the L2 span was significantly longer for high-proficiency bilinguals than for low-proficiency bilinguals, there was only a nonsignificant trend for the L1 span. More powerful designs will have to be used to find out whether this trend is a genuine effect, as suggested by Papagno and Vallar (1995) or a difference too small to be of practical value.

The issue whether the working memory capacity of bilinguals is larger than that of monolinguals has implications for the literature about the extent to which this capacity can be increased by training or is stable within an individual (Shipstead, Redick, & Engle, 2010). The question whether or not working memory can be trained is currently highly contested (Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Shipstead, et al., 2010). Some findings suggest that training one cognitive ability may extend to other cognitive abilities and, hence, be beneficial in general. Jaeggi, Buschkuehl, Jonides, and Perrig (2008), for example, reported higher fluid intelligence in participants who were trained on a *n*-back task, a working memory task which heavily relies on executive control resources. Other researchers, however, have identified several methodological concerns with these artificial training studies and claim that to this day no study has convincingly demonstrated that cognitive abilities can be trained, over and above (strategic) improvements in specific task demands (Shipstead, et al., 2010).

Further interesting findings of Vejnovic, et al. (2010) were that even for the highly proficient L2-speakers the reading spans were substantially shorter in L2 than in L1, and that there was quite high correlation between the L1 and L2 reading spans. The latter finding is in line with the assumption that reading spans measure stable individual differences in working memory capacity. The former finding agrees with the observation that L2 processing is more

demanding than L1 processing. As Hummel and French (2010) argued, the shorter working

memory span in L2 than L1 is likely to have implications for learning in L2, because it may

put bilinguals at the same disadvantage as monolinguals with a reduced memory span.

Hummel and French suggested that one way to counter this problem might be to provide L2

learners with additional written support when spoken instruction is used extensively (as in

communicative classroom contexts or immersion projects), so that L2 learners can try to

decrease the working memory load.

Conclusions

The aim of this chapter was to demonstrate that although the relation between memory and

language has remained poorly specified for many years, the last decade has witnessed an

important step forward in understanding how these key features of the human cognitive

system interact. On the one hand, native and foreign language acquisition appear to be

achieved through the ability to represent serial-order information in working memory, while

language perception and production rely on attentional control functions. The latter functions

are not specific to language processing and their continuous use seems to provide the

bilingual brain with a greater mental flexibility, although it may not increase working memory

capacity as measured by (reading) span tasks.

References

Abutalebi J., & Costa, A. (2008). Acquisition, Processing and Loss of L2: Functional, cognitive and neural perspectives. *Journal of Neurolinguistics, 21*, 473-476.

Ackerman, P.L., Beier, M.E., & Boyle, M.O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30-60.

Aron A.R, Watkins, L., Sahakian, B.J., Monsell, S., Barker, R.A., & Robbins, T.W. (2003). Task-set switching deficits in early-stage Huntington's disease: Implications for basal ganglia function. *Journal of Cognitive Neuroscience, 15,* 629-642.

Baddeley, A.D. (1993) Short-term phonological memory and long-term learning: A single-case study. *European Journal of Cognitive Psychology*, *5*, 129–148.

Baddeley, A.D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews: Neuroscience, 4,* 829-839.

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

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.H. Bower (Ed.) *The psychology of learning and motivation* Vol. 8 (47–89). New York: Academic Press.

Bialystok, E., Craik, F.I.M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*, 459–464.

Bialystok, E., Craik, F.I.M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19,* 290-303.

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified anti-saccade

    task: Effects of aging and bilingualism. *Journal of Experimental Psychology:*

    Learning, Memory, and Cognition, *32*, 1341–1354.

Binet, A., & Simon, Th. (1905). Application des méthodes nouvelles au diagnostic du niveau

    intellectual chez les enfants normaux et anormaux d'hospice et d'école primaire.

    L'Année Psychologique, 11, 245-336.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict

    monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Bowey, J. A. (2001) Nonword repetition and young children's receptive vocabulary: A

    longitudinal study. *Applied Psycholinguistics*, *22*, 441-469.

Brass, M., & von Cramon, D.Y. (2002). The role of the frontal cortex in task preparation.

    Cerebral Cortex, 12, 908-914.

Brass, M., & von Cramon, D.Y. (2004). Selection for cognitive control. An fMRI study on the

    selection of task relevant information. *The Journal of Neuroscience, 24*, 8847-8852.

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the revised hierarchical model

    of bilingual language processing after 15 years of service? *Bilingualism: Language*

    and Cognition, 13, 359-371.

Burgess, N. & Hitch, G. J. (1999). Memory for serial order: A network model of the

    phonological loop and its timing. *Psychological Review, 106,* 551-581.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term

    learning of verbal sequences. *Journal of Memory and Language, 55*, 627-652.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sen- tence comprehension.

    Behavioral and Brain Sciences, 22, 77–94.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary

learning. *Developmental Psychology, 32*, 867–873.

Costa, A. & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence

from language switching in highly proficient bilinguals and L2 learners. *Journal of

Memory and Language, 50*, 491-511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control

their lexicalization process? Inhibitory and Language-Specific Selection mechanisms

are both functional. *Journal of Experimental Psychology: Learning, Memory and

Cognition, 32*, 1057-1074.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their

mutual constraints within the human information-processing system. *Psychological

Bulletin, 104*, 163-191.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J.T., Aso, T.,

Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green D.W., & Price C.J . (2006).

Language control in the bilingual brain. *Science, 312,* 1537-1540.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and

reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Daneman, M., & Merikle, P.M. (1996). Working memory and language comprehension: A

meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433.

Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal working memory is

involved in associative word learning unless visual codes are available. *Journal of

Memory & Language, 48,* 527-541.

Emmorey, K., Luk, G., Pyers, J.E., & Bialystok, E. (2008). The source of enhanced cognitive

control in bilinguals. *Psychological Science, 19,* 1201-1206.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & 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.

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

Gathercole, S.E. (2006). Complexities and constraints in nonword repetition and word learning. *Applied Psycholinguistics, 27,* 599-613.

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

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

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

Goethe, K., & Oberauer, K. (2008). The integration of familiarity and recollection information in short-term recognition: Modeling speed-accuracy trade-off functions. *Psychological Research, 72*, 289-303.

Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 246-269.

Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81.

Gupta, P. (2005). Primacy and Recency in Nonword Repetition. *Memory, 13,* 318-324.

Gupta, P. (2009). A computational model of nonword repetition, immediate serial recall, and

    nonword learning. In A. Thorn and M. Page (Eds.), *Interactions between short-term*

    and long-term memory in the verbal domain. Hove, UK: Psychology Press.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P-H. (2005). Serial Position Effects in Nonword

    Repetition. *Journal of Memory and Language, 53*, 141-162.

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine

    vocabulary learning? Toward a computational resolution of the debate. *Journal of*

    memory and language, 61, 481-502.

Hebb, D.O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye

    (Ed.), *Brain mechanisms and learning* (pp. 37-46). Oxford: Blackwell.

Hernandez, A.E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching

    and language representation in Spanish-English bilinguals: An fMRI study.

    Neuroimage, 14, 510–520.

Hernandez, A.E., Martinez, A., & Kohnert, K. (2000). In search of the language switch: An

    fMRI study of picture naming in Spanish-English bilinguals. *Brain and Language, 73,*

    421–431.

Hummel, K., & French, L. (2010). Phonological memory and implications for the second

    language classroom. *Canadian Modern Language Review, 66*, 371-391.

Ivanova, I., & Costa, A. (2008) Does bilingualism hamper lexical access in speech

    production? *Acta Psychologica, 127,* 277–288.

Jacobs, J. (1887). Experiments on "prehension." *Mind, 12*, 75-79.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid

    intelligence with training on working memory. *Proceedings of the National Academy*

    of Sciences of the United States of America, 105, 6829-6833.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short and long-term benefits of

cognitive training. *Proceedings of the National Academy of Sciences of the United

States of America, 108,* 10081–10086.

Jarrold, C., & Towse, J.N. (2006). Individual differences in working memory. *Neuroscience,

139*, 39-50.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehen sion: Individual

differences in working memory. *Psychological Review, 98,* 122–149.

Kennedy A, Brooks R, Flynn L-A, & Prophet C. (2009). The reader's spatial code. In: Hyönä

J, Radach R. &, Deubel H. (Eds.) *The mind's eye: Cognitive and applied aspects of eye

movement research*. North-Holland: Amsterdam,pp. 193–212.

Lauro, L.J., Reis, J., Cohen, L.G., Cecchetto, C., & Papagno, C. (2010). A case for the

involvement of phonological loop in sentence comprehension. *Neuropsychologia*, 48,

4003–4011.

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A.L., Martinez, T., Schmidt, C., Weekes,

B., & Maquet , P. (2008). Neural networks for short-term memory for order

differentiate high and low proficiency bilinguals. *NeuroImage, 42,* 1698-1713.

Majerus, S., Lekeu, F., Van der Linden, M., & Salmon, E. (2001). Deep dysphasia: Further

evidence on the relation- ship between phonological short-term memory and lan-

guage processing impairments. *Cognitive Neuropsychology, 18,* 385–410.

Majerus, S., Poncelet, M., Elsen, B., & Van der Linden, M. (2006). Exploring the relationship

between new word learning and short-term memory for serial order recall, item recall,

and item recognition. *European Journal of Cognitive Psychology, 18,* 848-873.

Majerus, S., Poncelet, M., Van der Linden, M., & Weekes, B. (2008). Lexical learning in

    bilingual adults: the relative importance of short-term memory for serial order and

    phonological knowledge. *Cognition,107,* 395-419.

Majerus, S., Van der Linden, M., Mulder, G., Meulemans, T., & Peters, F. (2004). Verbal

    short-term memory reflects the sublexical organization of the phonological language

    network: Evidence from an incidental phonotactic learning paradigm. *Journal of

    Memory and Language, 51,* 297-306.

Martin, N. & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence

    from a single case study. *Brain and Language, 43*, 240-274.

Meuter, R. F. I. & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical

    costs of language selection. *Journal of Memory and Language, 40*, 25–40.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our

    capacity for processing information. *Psychological Review, 63*, 81-97.

Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the

    acquisition of phonological forms in children. *The Quarterly Journal of Experimental

    Psychology, 61*, 505-514.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and

    lexical, narrative, and grammatical skills in second-language oral production by adult

    learners. *Applied Psycholinguistics, 27,* 377-402.

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation,

    51*, 45-100.

Oztekin, I., & McElree, B. (2007). Proactive interference slows recognition by eliminating

    fast assessments of familiarity. *Journal of Memory and Language, 57*, 126-149.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial

recall. *Psychological Review, 105,* 761-781.

Page, M.P.A. & Norris, D. (2009). A Model Linking Immediate Serial Recall, the Hebb

Repetition Effect and the Learning of Phonological Word-forms. *Philosophical*

Transactions of the Royal Society B, 364, 3737-3753.

Papagno, C. & Vallar, G. (1995). Short-term memory and vocabulary learning in polyglots.

Quarterly Journal of Experimental Psychology, 48A, 98-107.

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and

foreign-language learning. *Journal of Memory and Language*, *30*, 331–347.

Service, E. (1992). Phonology, working memory, and foreign- language learning. *Quarterly*

Journal of Experimental Psychology, 45A, 21–50.

Prior, A., & Gollan, T.H. (in press). Good language-switchers are good task switchers:

Evidence from Spanish-English and Mandarin-English Bilinguals. *Journal of the*

International Neuropsychological Society.

Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., & Nieuwenhuis, S. (2004). The role of

medial frontal cortex in cognitive control. *Science, 306,* 443-447.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training

generalize? *Psychologica Belgica, 50*, 245-276.

Swets, B., Desmet, T., Hambrick, D.Z., & Ferreira, F. (2007). "The role of working memory

in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental*

Psychology: General, 136, 64-81.

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011). Control of

interference during working memory updating. *Journal of Experimental Psychology:*

Human Perception and Performance, 37, 137-151.

Szmalec, A., Duyck, W., Vandierendonck, A., Barberá Mata, A., & Page, M.P.A. (2009). The

Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly

Journal of Experimental Psychology, 62,* 435-443.

Szmalec, A., Loncke, M., Page, M., & Duyck, W. (2011). Order or disorder? Impaired Hebb

learning in dyslexia. *Journal of Experimental Psychology: Learning, Memory, and

Cognition, 37,* 1270-1279.

Vallar, G. & Baddeley, A.D. (1984). Fractionation of working memory. Neuropsychological

evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal

Behaviour, 23,* 151–161.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism

change native-language reading? Cognate effects in sentence context. *Psychological

Science, 20*, 923-927.

Vandierendonck, A. & Szmalec, A. (Eds) (2011). *Spatial Working Memory*. Hove, UK:

Psychology Press.

Vejnović, D., Milin, P., & Zdravković, S. (2010). Effects of Proficiency and Age of

Language Acquisition on Working Memory Performance in Bilinguals.

Psihologija, 43, 219-232.

Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural basis of asymmetric

language switching in second-language learners: An ER-fMRI study. *Neuroimage, 35,*

862–870.

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique

of Just and Carpenter (1992). *Psychological Review, 103,* 761–772.

Whitney, C., & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of

Research in Reading, 28,* 274–301.