Available online at www.sciencedirect.com
ScienceDirect

Journal of Memory and Language 56 (2007) 246–269

Journal of Memory and Language

www.elsevier.com/locate/jml

The nature of working memory in linguistic, arithmetic and spatial integration processes

Evelina Fedorenko *, Edward Gibson, Douglas Rohde

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 8 January 2006; revision received 24 June 2006
Available online 8 September 2006

Abstract

This paper reports the results of four dual-task experiments that were designed to determine the extent of domain-specificity of the verbal working memory resources used in linguistic integrations. To address this question, syntactic complexity was crossed in a 2 × 2 design with the complexity of a secondary task, which involved either (1) arithmetic integration processes and therefore relied on the use of verbal working memory, or (2) spatial integration processes. Experiments 1 and 2 crossed syntactic complexity and arithmetic complexity, and each revealed two main effects and a super-additive interaction during the critical region of the linguistic materials. Experiments 3 and 4 crossed syntactic complexity and the complexity of a spatial integration task, which does not rely on verbal working memory resources. Similar to Experiments 1 and 2, there were two main effects, but in contrast to Experiments 1 and 2, no interaction was observed in either experiment. The results of the four experiments show that linguistic processing interacts on-line with tasks that involve arithmetic but not spatial integration processes, suggesting that linguistic processing and other verbal working memory tasks that involve similar integration processes rely on a shared pool of working memory resources.

Keywords: Working memory; Language processing; Domain-specificity

The nature of working memory resources used in language processing has long been debated. Whereas there is ample evidence in favor of the independence of working memory resource pools used in verbal vs. visuo-spatial processing (e.g., Baddeley, 1986; Hanley, Young, & Pearson, 1991; Jonides et al., 1993; Logie, 1986, 1995; Paulesu, Frith, & Frackowiak, 1993; Shah & Miyake, 1996; Vallar & Shallice, 1990), it is less clear whether on-line language processing relies on the general verbal working memory resource pool (Fedorenko, Gibson, & Rohde, 2006; Gordon, Hendrick, & Levine, 2002; Just & Carpenter, 1992; King & Just, 1991), or whether it relies on a domain-specific resource sub-pool within the verbal working memory resource pool (Caplan & Waters, 1999).

Two approaches have traditionally been used to address the question of working memory resources used in on-line language processing: (1) an individual-differences approach, and (2) a dual-task approach. In the individual-differences approach, participants are divided into two or more groups on the basis of their performance on some form of a verbal working memory task (usually a reading span task (Daneman & Carpenter, 1980)) and tested on linguistic structures of varying syntactic complexity. In the dual-task approach, on the other hand, participants perform two tasks simultaneously: (1) on-line sentence processing, and (2) a

* Corresponding author.
E-mail address: evelina9@mit.edu (E. Fedorenko).

0749-596X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2006.06.007

non-linguistic verbally mediated task (usually a digit-/word-span task). The underlying assumption of the two approaches is that syntactic complexity will interact with group-type or with the difficulty of the secondary task, respectively, only if the non-linguistic verbally mediated memory task and on-line linguistic processing rely on overlapping pools of verbal working memory resources.

King and Just (1991) and Just and Carpenter (1992) claimed to have provided some evidence in support of the hypothesis whereby language processing relies on the general pool of verbal working memory resources. This evidence consisted of differential behavior of low- and high-span readers, classified using Daneman and Carpenter's (1980) reading span task, in the processing of syntactic structures of low and high complexity (subject- vs. object-extracted relative clauses). However, Caplan and Waters (1999) could not replicate these findings in a series of studies.

Moreover, Waters, Caplan, and Rochon (1995) used the dual-task approach crossing syntactic complexity and the complexity of a digit-span task. Specifically, participants were asked to perform a sentence–picture matching task with and without concurrent verbal load which involved maintaining a string of digits (equal to the subject's span, or equal to subject's span minus one). The sentences were all semantically reversible and varied in syntactic complexity (involving subject- vs. object-extracted relative clauses). Waters et al. observed a main effect of concurrent task in the accuracy scores for the sentence–picture matching task, but no effect of syntactic complexity, and crucially, no interaction between the difficulty of concurrent task and syntactic complexity. Furthermore, Waters and Caplan (1999) replicated these findings using an enactment task instead of the sentence–picture matching task.[1] On the basis of these results and on the basis of the data from the individual-differences studies, Caplan and Waters (1999) argued for a specialized pool of verbal working memory resources used for on-line language processing.

It is possible, however, that the reason that the previous attempts to find an interaction between linguistic complexity and non-linguistic verbally mediated tasks failed is that the cognitive processes involved in the language-processing task and in the digit-/word-span task are qualitatively different. Specifically, the digit-/word-span task involves storing a string of digits or unrelated words. In contrast, the language-processing task involves integrating each incoming word into the evolving structural representation, updating this representation, then integrating the next word, and so on. It is therefore plausible that the digit-/word-span task and linguistic integrations rely on independent pools of working memory resources.

In most of the previous dual-task experiments, the standard complexity manipulation in the digit-/word-span task has involved varying the number of elements that have to be remembered (cf. Fedorenko et al., 2006; Gordon et al., 2002; to be discussed below). In contrast, the standard complexity manipulation in the language-processing task in the dual-task experiments has usually involved the contrast between subject-extracted/object-extracted relative clauses. The difference between subject- and object-extracted relative clauses is plausibly related to the difference in integration lengths in the relative clause region (Ford, 1983; Gibson, 1998, 2000; Gordon, Hendrick, & Johnson, 2001, 2004; Grodner & Gibson, 2005; cf. King & Just, 1991; McElree, Foraker, & Dyer, 2003). Specifically, in subject-extracted relative clauses (1a), the embedded verb "frustrated" is integrated locally to the immediately preceding pronoun "who" co-indexed with the noun phrase "the janitor". In contrast, in object-extracted relative clauses (1b), at the point of processing the embedded verb "frustrated" it is necessary to retrieve the noun phrase "the janitor" from memory to interpret it as the object of "frustrated", since it occurs earlier in the input.

1a. Subject-extracted: The janitor **who frustrated the plumber** lost the key on the street.
1b. Object-extracted: The janitor **who the plumber frustrated** lost the key on the street.

The difficulty of retrieving the noun phrase "the janitor" at the point of processing the verb "frustrated" in the object-extracted relative clause in (1b) compared to a local integration between the relative pronoun co-indexed with "the janitor" and "frustrated" in the subject-extracted relative clause in (1a) might be due to either (1) the passive decay of the representation of the noun phrase "the janitor" over time, or (2) interference of the intervening noun phrase "the plumber". There is evidence for both of these factors contributing to the difficulty of processing object-extracted relative clauses and other long-distance dependencies (e.g., Gordon et al., 2001, Gordon, Hendrick, & Johnson, 2004; Grodner & Gibson, 2005; Lewis, 1996; McElree et al., 2003; Van Dyke & Lewis, 2003).

[1] Caplan and Waters (1999) discuss the results from Waters, Caplan, and Hildebrandt (1987) where participants were asked to make timed plausibility judgments about sentences containing subject- and object-extracted relative clauses, with and without a concurrent digit-span task, and the reaction time and the accuracy data revealed a main effect of the digit-span task, a main effect of syntactic complexity and an interaction between the two factors in the participants, but not in the items analysis. Caplan and Waters do not discuss the implications of this specific set of results for their claim that on-line language processing relies on a specialized pool of verbal working memory resources.

Recently, Gordon et al. (2002) applied the idea of interference as the underlying cause of difficulty in object-extracted relative clauses (compared to subject-extracted relative clauses) to the question of the extent of domain-specificity of the verbal working memory resources in language processing using a dual-task approach. Gordon et al. used a novel complexity manipulation in the word-span task, varying the degree of similarity between the to-be-remembered words and the words in the sentence materials, and observed a significant interaction in response accuracy data, such that the similarity between the to-be-remembered words and the words in the sentence materials affected object-extracted relative clauses to a larger extent. Fedorenko et al. (2006) extended these results to on-line language processing using a similar manipulation. These results are inconsistent with a hypothesis whereby language processing relies on an independent pool of verbal working memory resources.

In addition to the evidence from Gordon et al.'s and Fedorenko et al.'s studies, recent evidence from the language-processing literature began to suggest that even the working memory resource pool used in language processing may not be homogeneous. Specifically, at least two different types of working memory costs have been proposed. One type of working memory cost has been argued to be associated with keeping track of incomplete syntactic dependencies (e.g., Chomsky & Miller, 1963; Chen, Gibson, & Wolf, 2005; Gibson, 1991, 1998, 2000; Lewis, 1996; Wanner & Maratsos, 1978). Another type of working memory cost has been argued to be associated with integrating incoming words to earlier positions in the sentence (e.g., Gibson, 1998, 2000; Gordon et al., 2001; Grodner & Gibson, 2005; Warren & Gibson, 2002; cf. Konieczny, 2000; Vasishth, 2003). In addition to behavioral evidence in support of these two types of working memory costs, there exists ERP evidence in support of each: working memory cost associated with processing incomplete syntactic dependencies (King & Kutas, 1995; Kluender & Kutas, 1993), and working memory cost associated with processing long-distance integrations (Kaan, Harris, Gibson, & Holcomb, 2002; Phillips, Kazanina, & Abada, 2005). Specifically, it was shown that processing incomplete syntactic dependencies is associated with a sustained left-lateralized anterior negativity, whereas processing long-distance integrations is associated with a late positivity. Furthermore, Fiebach, Schlesewsky, and Friederici (2002) and Felser, Clahsen, and Munte (2003) used event-related potentials to directly investigate the relationship between these two types of working memory costs in the same experiment and provided evidence for their independence. Given the evidence for two different types of working memory costs involved in language processing, it is possible that the relationship between the working memory system involved in language processing and other working memory systems is more complex than previously thought. Specifically, it is possible that there are two pools of working memory resources used in on-line language processing, and they differ in the extent of their domain-specificity and in the extent of their overlap with other working memory systems. This would imply that in investigating the nature of working memory resources in language processing, it might be necessary to consider the two different resource pools used in language processing separately.

This paper presents four dual-task experiments, which explore the relationship between the working memory resources involved in the integration processes in language processing and those involved in the integration processes in (1) non-linguistic verbally mediated tasks, and (2) spatial tasks. Specifically, we wanted to test whether a non-linguistic verbally mediated task which involves integration processes might interact with linguistic integrations due to some overlap in the nature of the integration processes involved. To address this question, we conducted two dual-task experiments which investigated the relationship between linguistic processing and arithmetic processing. Arithmetic additions are similar to linguistic integrations, such that in a series of consecutive additions, an incoming element—a number—is integrated into the current representation, resulting in an intermediate sum. The intermediate sum is then updated with the integration of each incoming number. If it is indeed the case that all cognitive processes involving integrations of verbal material are relying on the same/overlapping pools of working memory resources, then we should observe an interaction between the complexity of linguistic and arithmetic integrations.

To control for a possible confound in terms of domain-general attention-switching costs, which might contribute to the interactions observed in dual-task paradigms (as noted by Caplan & Waters, 1999), we conducted two additional dual-task experiments where we substituted the arithmetic task with a spatial integration task, which involves similar integration processes, but critically does not require the use of verbal working memory resources. The attention-switching account predicts an interaction, regardless of the nature of the secondary task, as long as the secondary tasks are matched for complexity across the experiments. In contrast, the shared working memory resource pool hypothesis predicts that the arithmetic tasks but not the spatial integration tasks should interact with the sentence processing task, because only arithmetic tasks rely on verbal working memory resources.

Before presenting our experimental results, it is important to acknowledge a possible limitation in interpreting the presence of super-additive interactions in dual-task experiments. Previous dual-task experiments in different areas of cognitive psychology, as well as the experiments reported in this paper, rely on the additive factors logic (Sternberg, 1969, following Donders, 1868–1869), as summarized, for example, by McClelland (1979): "the assumption that one experimental manipulation influences the duration of one stage and another manipulation influences the duration of another stage leads to the conclusion that the two factors will have additive effects on reaction time. On the other hand, factors that influence the duration of the same stage will generally interact with one another" (McClelland, 1979, p. 311). In the experiments presented in this paper the additive factors logic applies as follows. If one experimental manipulation (the difficulty of the language comprehension task) draws on one resource pool, and another experimental manipulation (the difficulty of the arithmetic or spatial integration task) draws on another resource pool, then reaction times should reveal strictly additive effects. If, however, the two experimental manipulations draw on the same/overlapping resource pools then reaction times should reveal super-additivity. A potential problem with this logic arises if reaction times increase super-linearly, which could result in a super-additive interaction even when the two experimental manipulations draw on different resource pools (cf. Loftus (1978) who discusses this issue with respect to cases where probability (e.g., accuracy) is the dependent measure). However, this issue is mitigated in the current experimental design because reaction time curves, unlike probability curves, do not tend to show super-linear trends (e.g., Sternberg, 1969).

Experiment 1

This experiment had a dual-task design, in which participants read sentences phrase-by-phrase, and at the same time were required to perform a series of simple additions. The on-line addition task is similar to on-line sentence comprehension in that an incoming element—a number—must be integrated into (i.e., added to) the representation constructed thus far: the working sum. Both tasks had two levels of complexity, resulting in a 2×2 design. Critically, there was no difference in linguistic complexity between the easy- and hard-arithmetic conditions: the complexity of the arithmetic task was manipulated in terms of the difficulty of the arithmetic integrations (by making the addends larger), while keeping the linguistic form of the two conditions identical (number plus number plus number, etc.). Therefore, if we observe a super-additive interaction between the two tasks when the complexity of both tasks is high, then we may infer that the working memory resources that are involved in performing the arithmetic task overlap with those that are involved in syntactic integration processes. In contrast, if language processing relies on an independent working memory resource pool, there should be no such interaction.

Methods

Participants

Forty-eight participants from MIT and the surrounding community were paid for their participation. All were native speakers of English and were naive as to the purposes of the study.

Design and materials

The experiment had a 2×2 design, crossing syntactic complexity (subject-extracted RCs, object-extracted RCs) with arithmetic complexity (simple additions (low initial addend, subsequent addends between 1 and 3) vs. complex additions (higher initial addend, subsequent addends between 4 and 6)).

The language materials consisted of 32 sets of sentences, having four different versions as in (2):

(2)
a. *Subject-extracted, version 1*:
The janitor | who frustrated the plumber | lost the key | on the street.
b. *Subject-extracted, version 2*:
The plumber | who frustrated the janitor | lost the key | on the street.
c. *Object-extracted, version 1*:
The janitor | who the plumber frustrated | lost the key | on the street.
d. *Object-extracted, version 2*:
The plumber | who the janitor frustrated | lost the key | on the street.

There were two levels of syntactic complexity—subject- and object-extractions—with four versions of each sentence in order to control for potential plausibility differences between the subject- and object-extracted versions of each sentence. As a result, no independent plausibility control is needed in this design. Each participant saw only one version of each sentence, following a Latin-Square design (see Appendix A for a complete list of linguistic materials).

The numbers for the addition task were randomly generated on-line for each participant with the following constraints: (1) the value of the initial addend in the easy-arithmetic condition varied from 1 to 10, whereas the value of the initial addend in the hard-arithmetic condition varied from 11 to 20, and (2) the addends varied from 1 to 3 in the easy-arithmetic condition and from 4 to 6 in the hard-arithmetic condition. There is evidence from the mathematical cognition literature (e.g., Ashcraft, 1992, 1995) showing that reaction times as well as error rates for performing addition operations increase as a function of the size of the addends. That was the motivation for our complexity manipulations.

In addition to the target sentences, 40 filler sentences with various syntactic structures other than relative clauses were included. The length and syntactic complexity of the filler sentences were similar to that of the target sentences. The stimuli were pseudo-randomized separately for each participant, with at least one filler separating the target sentences.

Procedure

The task was self-paced phrase-by-phrase reading with a moving-window display (Just, Carpenter, & Woolley, 1982). The experiment was run using the Linger 2.85 software by Doug Rohde. Each experimental sentence had four regions (as shown in (2a)–(2d)): (1) a noun phrase, (2) an RC (subject-/object-extracted), (3) a main verb with a direct object (an inanimate noun phrase), and (4) an adjunct prepositional phrase. The addends for the addition task were presented simultaneously with the sentence fragments, above and aligned with the second character of each fragment. The first sentence region had a number above it (e.g., "12") and all the subsequent regions had a plus sign followed by a number (e.g., "+4"), as shown in Fig. 1.

Each trial began with a series of dashes marking the length and position of the words in the sentence. Participants pressed the spacebar to reveal each region of the sentence. As each new region appeared, the preceding region disappeared along with the number above it. The amount of time the participant spent reading each region and performing the accompanying arithmetic task was recorded as the time between key-presses.

To make sure the participants performed the arithmetic task, a window appeared at the center of the screen at the end of each sentence and the participants were asked to type in the sum of their calculations. If the answer was correct, the word "CORRECT" flashed briefly on the screen, if the answer differed by up to 2 from the correct sum, the word "CLOSE" flashed briefly, and if the answer was off by more than 2, the word "INCORRECT" flashed briefly on the screen. To assure that the participants read the sentences for meaning, two true-or-false statements were presented sequentially after the sum question, asking about the propositional content of the sentence they just read. Participants pressed one of two keys to respond "true" or "false" to the statements. After a correct answer, the word "CORRECT" flashed briefly on the screen, and after an incorrect answer, the word "INCORRECT" flashed briefly.

Participants were instructed not to concentrate on one task (reading or additions) more than the other. They were asked to read sentences silently at a natural pace and to be sure that they understood what they read. They were also told to answer the arithmetic and sentence questions as quickly and accurately as they could, and to take wrong answers as an indication to be more careful.

Before the experiment started, a short list of practice items and questions was presented in order to familiarize the participants with the task. Participants took ≈35 min to complete the experiment.

Results

Arithmetic accuracy

Participants answered the arithmetic sum question correctly 89.5% of the time. Table 1 presents the mean arithmetic accuracies across the four conditions of Experiment 1. A two-factor ANOVA crossing arithmetic complexity (easy, hard) and syntactic complexity (easy, hard) on these data revealed a main effect of arithmetic complexity ($F1(1,47) = 7.87$; $MSe = 0.0941$;

Table 1
Arithmetic accuracies in percent correct, as a function of arithmetic complexity and syntactic complexity (standard errors in parentheses)

Arithmetic complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy arithmetic	93.5 (1.7)	89.8 (1.8)
Hard arithmetic	86.7 (2.4)	87.8 (2.0)

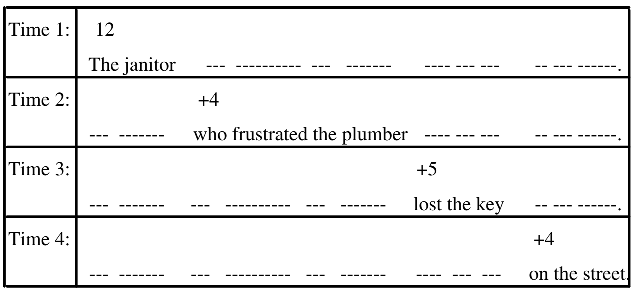

Fig. 1. Sample frame-by-frame presentation of an item in Experiment 1.

$p < .01$; $F2(1,31) = 8.12$; $MSe = 0.0627$; $p < .01$; min $F'(1,75) = 3.99$; $p < .05$), but no other reliable effects.

Comprehension question performance

There were two comprehension questions following each experimental trial. Participants answered the first question correctly 81.5% of the time, and the second question 79.4% of the time. The percentages of correct answers by condition were very similar for the two questions, so we collapsed the results in our analyses. Table 2 presents the mean accuracies across the four conditions of Experiment 1. A two-factor ANOVA crossing arithmetic complexity (easy, hard) and syntactic complexity (easy, hard) on the responses to the two comprehension questions revealed a main effect of syntactic complexity ($F1(1,47) = 13.37$; $MSe = .1270$; $p < .001$; $F2(1,31) = 6.41$; $MSe = .0846$; $p < .02$; min $F'(1,59) = 4.33$; $p < .05$) and a main effect of arithmetic complexity in the participants analysis ($F1(1,47) = 6.08$; $MSe = .0661$; $p < .02$; $F2(1,31) = 3.52$; $MSe = .0441$; $p = .07$; min $F'(1,63) = 2.22$; $p = .14$), but no significant interaction (Fs < 1.5).

Reaction times

Because participants had to answer three questions (one arithmetic, two language) for each sentence, the odds of getting all three correct were not very high overall (57.9%). As a result, we analyzed all trials, regardless of how the arithmetic and the comprehension questions were answered. The data patterns were very similar in analyses of smaller amounts of data, in which we analyzed (1) trials in which one or both of the comprehension questions were answered correctly, or (2) trials in which the arithmetic question was answered correctly. To adjust for differences in region lengths as well as overall differences in participants' reading rates, a regression equation predicting reaction times from region length was derived for each participant, using all filler and target items (Ferreira & Clifton, 1986; see Trueswell, Tanenhaus, & Garnsey, 1994, for discussion). For each region, the reaction time predicted by the participant's regression equation was subtracted from the actual measured reaction time to obtain a residual reaction time. Reaction time data points that were more than three standard deviations away from the mean residual RT for a position within a condition were excluded from the analysis, affecting 2.3% of the data. Fig. 2 presents the mean residual RTs per region across the four conditions of Experiment 1.

We conducted a 2 × 2 ANOVA crossing syntactic complexity and the complexity of the arithmetic task for each of the four regions. The results are presented in Table 3. For comparisons between means of conditions, we report 95% confidence intervals (CIs) based on the mean squared errors of the relevant effects from the participants analyses (see Masson & Loftus, 2003). We first present the analysis of the critical region, Region 2, which included the relative clause ("who frustrated the plumber"/"who the plumber frustrated"), followed by the analyses of the other regions. At the critical region, the hard-arithmetic conditions were read significantly slower than the easy-arithmetic conditions (380.8 ms vs. −49.5 ms; 95% CI = 120.6 ms), and the syntactically more complex object-extracted RC conditions were read significantly slower than the subject-extracted conditions (387.8 ms vs. −56.5 ms; 95% CI = 147 ms). Most interestingly, there was a significant

Table 2
Comprehension accuracies in percent correct, as a function of arithmetic complexity and syntactic complexity (standard errors in parentheses)

Arithmetic complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy arithmetic	85.8 (2.1)	78.8 (2.7)
Hard arithmetic	80.2 (2.4)	77.0 (2.1)

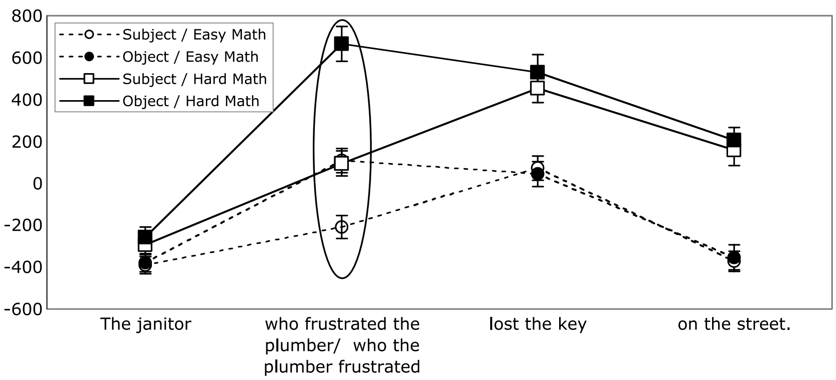

Fig. 2. Reaction times per region in the four conditions of Experiment 1. Error bars indicate standard errors. The critical region is circled.

Table 3
Analysis of variance results for Experiment 1

Source of variance	By participants			By items		min F'	
	df	$F1$ value	*MSe*	*df*	$F2$ value	*df*	min F'
Region 1							
Syntactic complexity	1,47	<1	39,434	1,31	<1	1,64	<1
Arithmetic complexity	1,47	10.74*	52,455	1,31	7.07*	1,66	4.26*
Interaction	1,47	<1	58,064	1,31	<1	1,63	<1
Region 2 (critical region containing the relative clause)							
Syntactic complexity	1,47	36.96*	256,350	1,31	43.46*	1,77	19.9*
Arithmetic complexity	1,47	51.53*	172,459	1,31	41.67*	1,71	23.0*
Interaction	1,47	4.40*	171,584	1,31	5.78*	1,78	2.49
Region 3							
Syntactic complexity	1,47	<1	149,644	1,31	<1	1,77	<1
Arithmetic complexity	1,47	40.78*	220,174	1,31	73.36*	1,77	26.2*
Interaction	1,47	<1	141,423	1,31	<1	1,76	<1
Region 4							
Syntactic complexity	1,47	<1	102,210	1,31	<1	1,78	<1
Arithmetic complexity	1,47	86.75*	164,414	1,31	152.81*	1,78	55.3*
Interaction	1,47	<1	85,723	1,31	<1	1,39	<1

Note. Significant effects are marked by asterisk.

interaction, such that in the hard-arithmetic conditions, the difference between subject- and object-extracted RCs (569.7 ms) was larger than in the easy-arithmetic conditions (318.8 ms). The statistical analyses of the raw reaction time data produced the same numerical patterns: specifically, the two main effects were significant in the participants and in the items analyses; and the interaction was significant in the participants analysis, but did not quite reach significance in the items analysis ($p = .12$). The interaction between syntactic and arithmetic complexity is predicted by the hypothesis whereby linguistic processing and arithmetic processing rely on overlapping pools of working memory resources, but not by the hypothesis whereby the pools of resources are independent.

In the other three regions (Region 1, Region 3, and Region 4) there was a main effect of arithmetic complexity, but no other significant effects.

Discussion

The results of Experiment 1 provide support for a working memory framework where linguistic integrations and arithmetic integrations rely on overlapping resource pools. Most importantly, there was an interaction between syntactic complexity and arithmetic complexity in the critical region of the linguistic materials, where syntactic complexity was manipulated between subject-extracted RCs (easy integrations) and object-extracted RCs (more difficult integrations). There was no evidence of any interaction of this kind in any of the other three regions. Critically, linguistic complexity did not vary across the two conditions of the arithmetic task (both conditions used expressions like “number plus number plus number”, etc.), so the observed interaction is not due to an overlap in the linguistic processes that are involved in the two tasks. In other words, the fact that the arithmetic task uses verbal material cannot, by itself, account for the observed interaction.

It should be noted, however, that there are two possible confounding factors present in Experiment 1. The first confounding factor involves a difference between the easy and the hard conditions of the arithmetic task in terms of low-level verbal complexity, and the second confounding factor involves a possible explanation of the interaction in terms of a domain-general attention-switching mechanism.

First, even though the easy and the hard conditions in the arithmetic task are the same in terms of syntactic complexity, there might be a difference between the two conditions in terms of low-level morpho-/phonological complexity, which might result in the hard conditions having higher rehearsal demands. Specifically, because the hard conditions involve adding larger numbers, both the length and the morphological complexity of the numbers in the hard conditions are on average higher. In the working memory framework proposed by Baddeley and colleagues (Baddeley, 1986; Baddeley & Hitch, 1974), this would involve the phonological loop system used for storing and rehearsing verbal material. On the assumption that both the linguistic and the arithmetic task make use of the phonological loop, one could argue that the observed interaction might be due to the fact that the hard language conditions (object-extracted relative clauses) and the hard-arithmetic conditions (more difficult additions) have higher rehearsal demands and are thus overtaxing the rehearsal system. To rule out the possibility that an overlap in the rehearsal system use between the linguistic and the arithmetic task contributed to the interaction observed in Experiment 1, we conducted Experiment 2—where the morpho-/phonological complexity was kept constant across the easy and the hard conditions of the arithmetic task and only the difficulty of the arithmetic operations was manipulated. (It is worth noting, however, that the explanation in terms of overtaxing the shared rehearsal system is not very plausible, given the patterns of data in the previous experiments in the literature. Specifically, as discussed above, in the earlier dual-task experiments where a digit-span or a word-span task was used as a secondary task, the complexity manipulation (more vs. fewer items to remember) inevitably varied the amount of required rehearsal, and yet no reliable interactions between digit-/word-span task complexity and syntactic complexity have been observed (Caplan & Waters, 1999; Waters et al., 1995).)

Another alternative explanation for the observed pattern of results in Experiment 1 is in terms of attentional resources required for the simultaneous performance of the two tasks, as discussed in Caplan and Waters (1999). In dual-task paradigms, resources are needed in order to direct attention to one task or another. It is possible that in the difficult conditions, more attention switches are required, or the switches between tasks are more costly. The observed interaction could therefore be a result of additional task-switching costs in the high syntactic complexity/high arithmetic complexity condition. Experiments 3 and 4—where an arithmetic task was substituted by a spatial integration task—were conducted to address this issue.

Experiment 2

This experiment had a similar dual-task design, in which participants read sentences phrase-by-phrase, and at the same time were required to perform arithmetic calculations. In contrast to Experiment 1, the difficulty of the arithmetic task was manipulated by using different operations in the easy and hard conditions: additions were used in the easy conditions and subtractions were used in the hard conditions. There is evidence that subtractions take longer and are more error prone than additions (Campbell & Xue, 2001) and also that children learn additions before subtractions (e.g., Siegler, 1987), suggesting that there is something more difficult about the process of subtraction, compared to addition. The range of numbers used in both conditions was the same. Therefore, if we observe a super-additive interaction between syntactic integrations and the new arithmetic task, then we may infer that the interaction observed in Experiment 1 was not due to the difference in rehearsal demands between the easy and hard conditions of the arithmetic task, and that the working memory resources that are involved in performing the arithmetic task overlap with those that are involved in syntactic integration processes. In contrast, if language processing relies on an independent working memory resource pool, there should be no such interaction.

Methods

Participants

Forty participants from MIT and the surrounding community were paid for their participation. All were native speakers of English and were naive as to the purposes of the study. None participated in Experiment 1.

Design and materials

The experiment had a 2×2 design, crossing syntactic complexity (subject-extracted RCs, object-extracted RCs) with arithmetic complexity (simple arithmetic operations (initial addend between 30 and 50, subsequent addends between 3 and 6, additions) vs. complex arithmetic operations (initial addend between 40 and 60, subsequent addends between 3 and 6, subtractions)). The size of the initial addend differed between the easy and hard conditions, so that across the regions participants have to work with the numbers in approximately the same range.

The language materials, including 40 fillers, were the same as those used in Experiment 1. The numbers for the arithmetic task were randomly generated on-line for each participant with the constraints described above. For the filler sentences, the arithmetic task had the following constraints: (1) the initial addend was between 30 and 60, and (2) the subsequent addends (with the values between 3 and 6) could be either added or subtracted.

Procedure

The procedure was exactly the same as in Experiment 1. Participants took ≈35 min to complete the experiment.

Results

Arithmetic accuracy

Participants answered the arithmetic sum correctly 85% of the time. Table 4 presents the mean arithmetic accuracies across the four conditions of Experiment 2. A two-factor ANOVA crossing arithmetic complexity (easy, hard) and syntactic complexity (easy, hard) on these data revealed no significant effects and no

Table 4
Arithmetic accuracies in percent correct, as a function of arithmetic complexity and syntactic complexity (standard errors in parentheses)

Arithmetic complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy arithmetic	85.6 (2.4)	85.3 (2.3)
Hard arithmetic	82.5 (2.4)	85.6 (2.2)

interaction (all $Fs < 1.5$). Notice that this pattern of results differs from that in Experiment 1 where a main effect of arithmetic complexity was observed.

Comprehension question performance

There were two comprehension questions following each experimental trial. Participants answered the first question correctly 78.5% of the time, and the second question 74.5% of the time. As in Experiment 1, we collapsed the results in our analyses. Table 5 presents the mean accuracies across the four conditions of Experiment 2. A two-factor ANOVA crossing arithmetic complexity (easy, hard) and syntactic complexity (easy, hard) on the responses to the two comprehension questions revealed no main effects and no interaction (all $Fs < 2$).

Table 5
Comprehension accuracies in percent correct, as a function of arithmetic complexity and syntactic complexity (standard errors in parentheses)

Arithmetic complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy arithmetic	78.6 (2.5)	74.4 (2.4)
Hard arithmetic	76.1 (2.3)	76.9 (2.6)

Note that this pattern of results is different from that in Experiment 1, where an effect of syntactic complexity and an effect of arithmetic complexity (significant in the participants analysis) were observed.

Reaction times

As in Experiment 1, we analyzed all trials, regardless of how the arithmetic and the comprehension questions were answered. Also, as in Experiment 1, reaction time data points that were more than three standard deviations away from the mean residual RT for a position within a condition were excluded from the analysis, affecting 1.9% of the data. Fig. 3 presents the mean residual RTs per region across the four conditions of Experiment 2.

We conducted a 2×2 ANOVA crossing syntactic complexity and the complexity of the arithmetic task for each of the four regions. The results are presented in Table 6. We first present the analysis of the critical region, Region 2, which included the relative clause ("who frustrated the plumber"/"who the plumber frustrated"), followed by the analyses of the other three regions. At the critical region, the hard-arithmetic conditions were read slower than the easy-arithmetic conditions (137.8 ms vs. −0.43 ms; 95% CI = 139.4 ms), and the syntactically more complex object-extracted RC conditions were read significantly slower than the subject-extracted conditions (234 ms vs. −96.7 ms; 95% CI = 159.5 ms). Most importantly, there was a significant interaction, such that in the hard-arithmetic conditions, the difference between subject- and object-extracted RCs was larger (546.4 ms) than in the easy-arithmetic conditions (114.9 ms). This interaction is predicted by the hypothesis whereby sentence processing and arithmetic processing rely on overlapping pools of WM resources, but not by the hypothesis whereby the pools of resources are independent.

In the other three regions, the patterns of reaction times were as follows. In Region 1, there was an unpre-

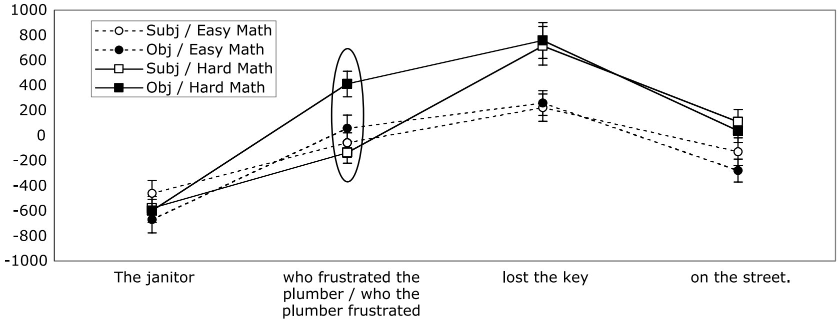

Fig. 3. Reaction times per region in the four conditions of Experiment 2. Error bars indicate standard errors. The critical region is circled.

Table 6
Analysis of variance results for Experiment 2

Source of variance	By participants			By items		min F'	
	df	$F1$ value	*MSe*	*df*	$F2$ value	*df*	min F'
Region 1							
Syntactic complexity	1,39	5.29*	98,456	1,31	4.88*	1,68	2.53
Arithmetic complexity	1,39	<1	99,559	1,31	<1	1,60	<1
Interaction	1,39	2.93	120,448	1,31	5.60*	1,67	1.92
Region 2 (critical region containing the relative clause)							
Syntactic complexity	1,39	17.58*	248,728	1,31	15.76*	1,68	8.31*
Arithmetic complexity	1,39	4.02	189,919	1,31	2.70	1,64	1.61
Interaction	1,39	12.03*	154,865	1,31	5.49*	1,56	3.76
Region 3							
Syntactic complexity	1,39	<1	414,711	1,31	<1	1,45	<1
Arithmetic complexity	1,39	13.86*	714,308	1,31	31.09*	1,65	9.58*
Interaction	1,39	<1	275,918	1,31	<1	1,41	<1
Region 4							
Syntactic complexity	1,39	1.79	271,729	1,31	2.39	1,70	1.02
Arithmetic complexity	1,39	15.11*	202,847	1,31	12.32*	1,67	6.78*
Interaction	1,39	<1	137,942	1,31	<1	1,44	<1

Note. Significant effects are marked by asterisk.

dicted effect of syntactic complexity, such that the object-extracted RCs (−633.3 ms) were read faster than the subject-extracted RC (−519.2 ms). There is no reason to expect a difference between subject- and object-extracted RCs in this region, because the linguistic materials were exactly the same. Similarly, there was an unpredicted interaction, such that the difference between the easy- and the hard-arithmetic conditions was larger in the subject-extracted conditions than in the object-extracted conditions. Again, there is no reason to expect any differences among the four conditions in this region, because the linguistic materials were exactly the same. In Regions 3 and 4, there was a main effect of arithmetic complexity, but no other effects and no interaction.

Discussion

In Experiment 2 we controlled for potential differences in rehearsal demands between the easy and hard versions of the arithmetic task, and we observed an interaction similar to that in Experiment 1. Specifically, a condition where both linguistic integrations were hard (object-extracted RCs) and arithmetic integrations were hard (subtractions) was especially difficult for subjects to perform, more than would be expected if the two complexity effects were additive.

The results of Experiment 2 allow us to rule out the explanation of the interaction observed in Experiment 1 in terms of the difference in rehearsal demands between the easy- and the hard-arithmetic conditions. Specifically, unlike in Experiment 1, in Experiment 2 the numbers used in the easy- and the hard-arithmetic conditions did not differ in terms of length and/or morphological complexity; only the difficulty of the operations themselves was manipulated. Despite this fact, we still observed an interaction between syntactic and arithmetic complexity in the critical region, such that when both tasks were difficult, participants experienced more difficulty than would be expected if the two effects were merely additive. Thus, the results of Experiment 2 provide further support for a working memory framework where language processing and arithmetic processing rely on overlapping WM resource pools.

However, as discussed above, there is another confound present in both Experiments 1 and 2. Specifically, it is possible to account for the observed interactions in terms of attention-switching costs: it is possible that in the hard conditions, more switches between the tasks are required, or the switches are more costly, regardless of the nature of the task. To address this issue, Experiments 3 and 4 were conducted. As discussed in the Introduction, ample evidence exists showing that different pools of working memory resources are used for verbal vs. visuo-spatial processing (e.g., Baddeley, 1986; Baddeley & Hitch, 1974; Hanley et al., 1991; Jonides et al., 1993; Shah & Miyake, 1996; Vallar & Shallice, 1990). Therefore, the attention-switching account predicts that an interaction similar to those observed in Experiments 1 and 2 should be observed regardless of the nature of the two tasks involved, as long as they are matched for difficulty with the tasks used in Experiments 1 and 2. In Experiments 3 and 4, we used the same linguistic materials as in Experiments 1 and 2, but we substituted the arithmetic task with a spatial integration task. We used two different versions of a spatial integration task in the two new experiments.

Experiment 3

This experiment used a similar dual-task paradigm as the first two experiments. In contrast to Experiments 1 and 2, however, the secondary task was a spatial-rotation task. In this task, participants were instructed to visually imagine adding different-size sectors of a circle and to keep track of the angle subtended by the combined segments. The most natural way to solve this task is to mentally rotate each incoming sector until it abuts the estimated sum of the previous sectors, as shown in Fig. 4.

The on-line spatial-rotation task is similar to the addition task in that an incoming element—a sector—must be integrated into, or added to, the representation constructed thus far. Critically though, the spatial-rotation task does not rely on verbal working memory resources, and should not therefore interact with the sentence-processing task if the cause for the observed interaction in Experiments 1 and 2 is an overlap in the use of verbal working memory resources. However, if the attentional costs are responsible for the interaction, we should observe a similar interaction, regardless of the nature of the secondary task.

In order to draw conclusions of this sort, however, it is necessary to assure that the spatial integration task is of approximately the same difficulty as the arithmetic tasks used in the first two experiments. Specifically, one reason for why an interaction might be observed when a given task (let us call it the primary task, for ease of discussion) is paired with one secondary task, but not when it is paired with another secondary task could be that one secondary task is easier than the other secondary task. In this case, in the experiment where the secondary task is easier, resources might be abundant, and thus the results would not speak to the relationship between the pools of resources used to perform the two tasks. In other words, even if the two tasks rely on the same pool/overlapping pools of resources, it is possible that no super-additive interaction would be observed due to the fact that the pools do not get overtaxed even in the condition where both tasks are complex. To make sure that our arithmetic tasks and our spatial integration tasks were comparable in difficulty, we had an independent group of participants perform the arithmetic task from Experiment 1 and the spatial integration task from Experiment 3 in isolation and we analyzed the reaction times and accuracies in these two tasks (see Norming for Task Difficulty).

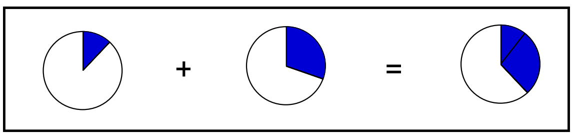

Fig. 4. Spatial-rotation task in Experiment 3.

Methods

Participants

Sixty-four participants from MIT and the surrounding community were paid for their participation. All were native speakers of English and were naive as to the purposes of the study. None participated in Experiments 1 or 2.

Design and materials

The experiment had a 2×2 design, crossing syntactic complexity (subject-/object-extracted RCs) with the complexity of the spatial-rotation task (simple rotations with small-angle sectors/complex rotations with larger-angle sectors). The language materials were the same as those used in Experiments 1 and 2.

The sectors for the spatial-rotation task were randomly generated on-line for each participant in the following way: the size of the sectors for the easy condition varied from 5° to 90°, whereas the size of the sectors for the hard condition varied from 30° to 180°. As a result, it was possible in the hard condition for the sum of sectors to be more than 360°, thus to "wrap around" the circle. Previous research (e.g., Shepard & Metzler, 1971) has shown that the time it takes subjects to rotate a two- or three-dimensional figure is related to the angle of the rotation, such that larger angles take longer. Furthermore, pilot testing of the pie task by itself suggested that the task is easier to perform with smaller sectors.

Procedure

The procedure was identical to that of Experiments 1 and 2, except for substituting the spatial-rotation task for the arithmetic tasks. Above each sentence fragment, participants saw a small circle. They were instructed to think of it as a plate for a pie. On each "plate", there was a "pie-slice" shown in blue. The size of the "pie-slices" varied (as described in Design and materials), but they all started at the vertically pointing radius position, as shown in Fig. 5.

Participants were instructed to visually imagine adding each new "pie-slice" to the previous one(s) by mentally "putting" them next to each other. To assure that the participants performed the task, at the end of each trial a large blank circle appeared at the center of the screen with a vertically pointing radius. Participants were instructed to drag this radius (using the mouse) to the end-point where all the "pie-slices" they just

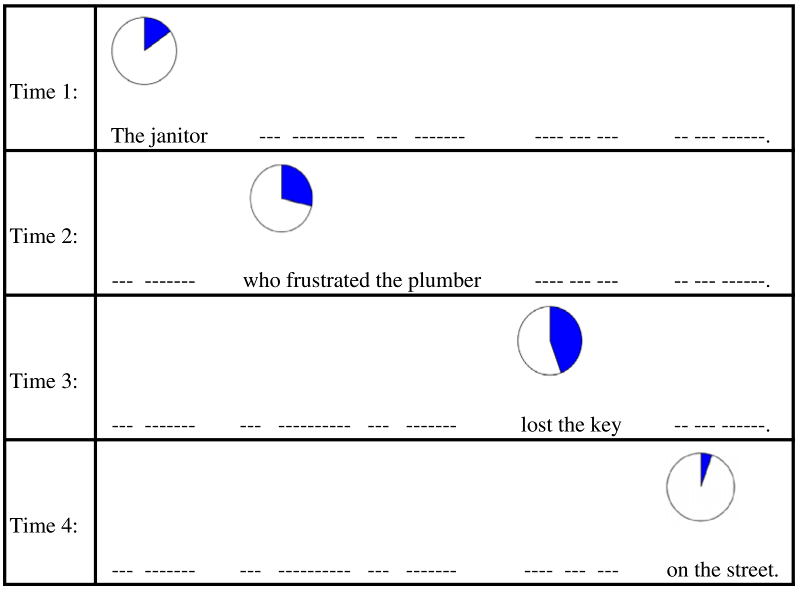

Fig. 5. Sample frame-by-frame presentation of an item in Experiment 3. (For interpretation of the references to colour in the text citation, the reader is referred to the web version of this paper.)

saw would come to when placed next to each other. If the answer was within 10° of the correct answer, the words "Very Close!" flashed briefly on the screen; if the answer was within 35°, the words "Pretty Good" flashed briefly; if the answer was within 90°, the words "In The Ballpark" flashed briefly; finally, if the answer was not within 90°, the words "Not Very Good" flashed briefly on the screen. The participants were warned that sometimes the "pie-slices", when added together, would form more than a complete pie. In such cases, they were told to assume that the slices "wrapped around" and to ignore the complete portion of the pie when keeping track of the end-point. As in Experiments 1 and 2, this task was followed by two comprehension questions about the content of the sentences.

Thus far, we have been referring to the task in Experiment 3 as a spatial integration task. In order to establish that participants were, in fact, performing this task using spatial working memory resources, and not verbal working memory resources, we administered a post-experimental questionnaire to try to understand the strategies the subjects might use in performing the pie task. The question about the strategies was open-ended, giving the subjects a chance to give any feedback they felt was relevant as to how they were performing the task. The open-ended nature of the question resulted in about half of the answers being impossible to code in terms of the strategy—spatial or verbal—used by the subject. The rest of the answers were coded as either "spatial-strategy" or "verbal-strategy": we marked the strategy as being "spatial" if the answer explicitly mentioned a spatial process, and we marked the strategy as being "verbal" if the answer explicitly mentioned a verbally mediated process, i.e., a process where verbal labels could be mapped onto the spatial chunks. Of the answers where the type of strategy could be identified (56% of the responses), 73% of the responses were coded as "spatial-strategy", and 27%—were coded as "verbal-strategy". Examples of spatial-strategy answers included things like "tried to visualize it", "I imagined the line rotating along with each piece", "would try to visualize as I went along", etc. Examples of verbal-strategy answers included things like "clock face patterns", "usually rounding the pie slices to easy chunks was helpful, i.e., if the slice looked almost like a quarter, I rounded it to a quarter", etc. Note that 27% is a conservative estimate, because even some of the answers, which were coded as "verbal-strategy" could, in principle, be performed in a spatial way: for example, quarters and halves are meaningful spatial chunks, and thus might be easier to operate on, compared to less meaningful sector-sizes. In other words, the fact that a participant would mention "quarters" does not necessarily imply the use of a verbal strategy. Thus, 27% represents an upper bound on the subjects that used a verbal strategy of those whose answers were codable. In addition to the questionnaire responses, several of the participants verbally reported that they initially tried a verbally based strategy (e.g., encoding the pie-slices in terms of the number of hours), but had to quickly switch to the spatial-rotation strategy, because they found it too difficult to perform the pie-task using a verbal strategy. Given this feedback, it seems safe to conclude that most of the participants performed the task via spatial rotation, as instructed, and thus relied on spatial, and not verbal, working memory resources. Furthermore, foreshadowing Experiment 4, it is worth noting that the spatial integration task in that experiment is much less subject to the criticism of potential reliance on verbal strategies, because unlike the spatial integration task in Experiment 3—where some verbal strategies seem possible—in Experiment 4 it seems difficult to devise a verbal strategy for solving the spatial integration task.

Norming for task difficulty

In the Norming Study, an independent group of 37 participants, none of whom participated in any of the four experiments described in this paper, were asked to perform two tasks in turn: (1) the arithmetic addition task from Experiment 1, and (2) the spatial integration task from Experiment 3. The order of tasks was counterbalanced across the participants.

In theory, it should be possible to compare two tasks in terms of their relative difficulty using two dependent measures—reaction times and accuracies. However, it is very difficult to use accuracies as a dependent measure of performance on these two tasks, because the answers are qualitatively very different, and it is difficult to compare them. Specifically, in the arithmetic task, participants provide an answer (a sum), which is either correct or incorrect, thus the accuracies are calculated as percent correct. In contrast, in the spatial integration task, participants are asked to drag the radius to the position subtended by all the sectors added together, as described above, and the accuracies are calculated as degrees off from the correct answer. There is no obvious way to map these two measures onto each other. Therefore, the primary dependent measure we use is reaction time. Moreover, because we are interested in how working memory resources are used in *on-line* processing, a reaction time measure is more informative. An anonymous reviewer has observed that it is difficult in general to meaningfully compare reaction times in two tasks that do not have comparable accuracy measures. Specifically, in order to argue that one task is more difficult than another, reaction times in the first task should be equal or longer, and the accuracies should be equal or lower. If the accuracies cannot be compared, then the possibility of speed-accuracy trade-offs arises, such that even though one task may take longer, it may be the case that participants are expending more effort to perform the task (and are therefore more accurate).

Whereas this issue can be a problem in comparing reaction times in some pairs of tasks, it is less relevant to the current comparison because the performance on the task which takes less time to perform (the arithmetic task) is at ceiling (97–99%). That is, (1) reaction times in the arithmetic task are faster than in the spatial-rotation task, and (2) the accuracies in the arithmetic task are at least as high as in the spatial-rotation task, because they are at ceiling. Thus, it is plausible that the spatial-rotation task is more difficult than the arithmetic task.

We first present a summary of the reaction time data and the accuracy data from the arithmetic task. Then, we present a similar summary for the spatial integration task. Finally, we present a direct comparison analysis for the two tasks in terms of reaction times.

In the arithmetic task, reaction times in every region revealed a significant effect of task complexity, such that more difficult additions took longer ($F\text{s} > 5$, $p\text{s} < .05$). Furthermore, the arithmetic task accuracies also revealed a significant effect of task complexity, such that more difficult additions were less accurate—97% vs. 99% ($F(1,36) = 10.2$; $p < .05$; $F2(1,31) = 4.57$; $p < .05$).

In the spatial integration task, similar to the arithmetic task, reaction times in every region revealed a significant effect of task complexity, such that more difficult rotations took longer ($F\text{s} > 5$, $p\text{s} < .05$). The spatial integration task accuracies also revealed a significant effect of task complexity, such that more difficult rotations were less accurate $-29°$ off from the correct answer vs. $23°$ ($F(1,36) = 17.9$; $p < .001$; $F2(1,31) = 15.5$; $p < .001$). Notice that the accuracies for both the arithmetic task and the spatial integration task were somewhat higher (although very comparable) in the Norming Study, compared to Experiments 1 and 3, respectively. This is expected given that the task demands are higher in the dual-task experiments, compared to the Norming Study where each task is performed in isolation.

We used paired-samples two-tailed *t*-tests to compare raw reaction times for the two tasks during (1) Region 2 (the critical region in the four experiments described in this paper), and (2) across all four regions. Both *t*-tests revealed that participants took longer to perform the spatial integration task. The average reaction times during the critical region were 1099 ms for the arithmetic task and 1867 ms for the spatial integration task ($t(1,36) = -7.85$, $p < .001$). The average reaction times during all the regions were 1125 ms for the arithmetic task and 1709 ms for the spatial integration task ($t(1,36) = -6.67$, $p < .001$). These results suggest that the spatial integration task was more difficult for participants to perform than the arithmetic task, when performed in isolation.

Results

Spatial-rotation task accuracy

On average, participants' estimates were 38.4° off from the correct answer. Table 7 presents the mean accuracies (in degrees off from the correct answer) across the four conditions of Experiment 3. A two-factor ANOVA crossing spatial-rotation task complexity (easy, hard) and syntactic complexity (easy, hard) revealed a main effect of

Table 7
Spatial-task accuracies in degrees off from the correct answer, as a function of spatial task complexity and syntactic complexity (standard errors in parentheses)

Spatial task complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy rotations	35.1 (2.9)	33.3 (2.4)
Hard rotations	42.9 (2.6)	42.4 (2.5)

complexity of the spatial-rotation task ($F1(1,63) = 19.31$; $MSe = 4621$; $p < .001$; $F2(1,31) = 25.63$; $MSe = 2295$; $p < .001$; min $F'(1,90) = 11.0$, $p < .002$), but no other significant effects (Fs < 1).

Comprehension question performance

There were two comprehension questions following each experimental trial. Participants answered the first question correctly 78.7% of the time, and the second question 77.8% of the time. As in the other experiments, we collapsed the results in our analyses. Table 8 presents the mean accuracies across the four conditions of Experiment 3. A two-factor ANOVA crossing spatial-rotation task complexity (easy, hard) and syntactic complexity (easy, hard) on the responses to the comprehension questions revealed a marginal effect of the spatial-rotation task complexity in the participants analysis ($F1(1,63) = 3.13$; $MSe = .0325$; $p = .082$; $F2 < 1$; min $F' < 1$), but no other effects or interactions (Fs < 1).

Table 8
Comprehension accuracies in percent correct, as a function of spatial task complexity and syntactic complexity (standard errors in parentheses)

Spatial task complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy rotations	79.7 (2.0)	79.0 (2.1)
Hard rotations	76.9 (2.0)	77.3 (1.9)

Reaction times

As in Experiments 1 and 2, we analyzed all trials, regardless of how the comprehension questions were answered and how the spatial-rotation task was performed. Also, as in Experiments 1 and 2, reaction time data points that were more than three standard deviations away from the mean residual RT for a position within a condition were excluded from the analyses, affecting 1.8% of the data. Fig. 6 presents the mean residual reaction times per region across the four conditions in Experiment 3.

We conducted a 2 × 2 ANOVA crossing syntactic complexity and the complexity of the spatial integration task for each of the four regions. The results are presented in Table 9. We first present the analysis of the critical region, Region 2, which included the relative clause ("who frustrated the plumber"/"who the plumber frustrated"), followed by the analyses of the other three regions. At the critical region, the hard-spatial-task conditions were read significantly slower than the easy-spatial-task conditions (293.8 ms vs. −133.9 ms; 95% CI = 158.2 ms), and the syntactically more complex object-extracted RC conditions were read significantly slower than the subject-extracted RC conditions (264.9 ms vs. −105.0 ms; 95% CI = 118.6 ms). Critically, there was no interaction between syntactic complexity and the complexity of the spatial task. Moreover, the effect of syntactic complexity in the hard-spatial-task conditions was numerically *smaller* (319.5 ms) than that in the easy-spatial-task conditions (420.3 ms). This result argues against the attentional explanation of the interaction that was observed in Experiments 1 and 2.

In Region 1, there were no significant effects. In Regions 3 and 4, there was a main effect of spatial task complexity, but no other significant effects.

Analysis of Experiment 1 and Experiment 3 with experiment as a factor

To further strengthen the conclusions we draw from the different patterns of results we observed in Experiment 1 (an interaction between linguistic and arithmetic

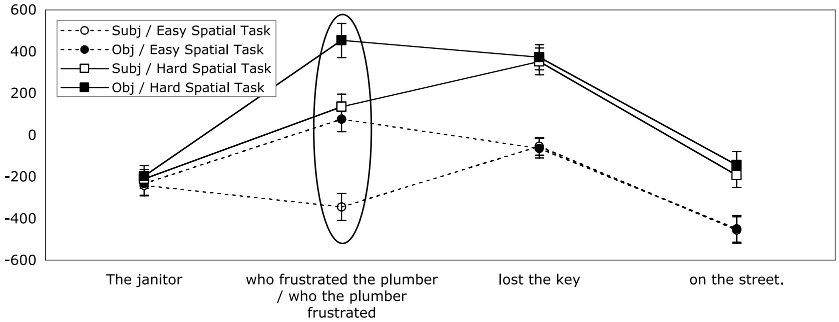

Fig. 6. Reaction times per region in the four conditions of Experiment 3. Error bars indicate standard errors. The critical region is circled.

Table 9
Analysis of variance results for Experiment 3

Source of variance	By participants			By items		min F'	
	df	$F1$ value	*MSe*	*df*	$F2$ value	*df*	min F'
Region 1							
Syntactic complexity	1,63	<1	76,995	1,31	<1	1,94	<1
Spatial-task complexity	1,63	<1	83,902	1,31	<1	1,69	<1
Interaction	1,63	<1	77,597	1,31	<1	1,52	<1
Region 2 (critical region containing the relative clause)							
Syntactic complexity	1,63	38.82*	225,555	1,31	47.77*	1,88	21.4*
Spatial-task complexity	1,63	29.19*	401,031	1,31	67.44*	1,94	20.3*
Interaction	1,63	1.10	147,629	1,31	1.38	1,89	<1
Region 3							
Syntactic complexity	1,63	<1	113,081	1,31	<1	1,93	<1
Spatial-task complexity	1,63	52.45*	217,636	1,31	88.86*	1,93	32.9*
Interaction	1,63	<1	138,094	1,31	<1	1,39	<1
Region 4							
Syntactic complexity	1,63	<1	69,797	1,31	<1	1,61	<1
Spatial-task complexity	1,63	30.78*	166,628	1,31	107.41*	1,89	23.9*
Interaction	1,63	<1	126,894	1,31	<1	1,91	<1

Note. Significant effects are marked by asterisk.

complexity) and in Experiment 3 (a lack of an interaction between linguistic and spatial integration task complexity), we analyzed the two datasets—from Experiments 1 and 3—using a $2 \times 2 \times 2$ ANOVA, with the following factors: (1) syntactic complexity (subject-extractions/object-extractions), (2) non-linguistic task (arithmetic/spatial-rotation) complexity (easy/hard), and (3) experiment (Experiment 1/Experiment 3). At the critical relative clause region we observed a significant three-way interaction, such that the interaction between syntactic and non-linguistic task complexity was observed only in Experiment 1, and not in Experiment 3 ($F1(1,110) = 4.84$; $MSe = 779539$; $p < .05$; $F2(1,31) = 12.2$; $MSe = 466202$; $p < .002$; min $F'(1,139) = 3.27$; $p = .07$). There was no such interaction in any of the other regions (Fs < 1).

We further examined the raw reaction times both at the critical region and across all the regions in Experiment 1 and Experiment 3, and we found that the time ranges were very similar. Specifically, across all the regions, the mean raw reaction time in Experiment 1 was 2187 ms ($SE = 43$ ms), and the mean raw reaction time in Experiment 3 was 2066 ms ($SE = 59$ ms); at the critical region, the mean raw RT in Experiment 1 was 2780 ms ($SE = 66$ ms), and the mean raw reaction time in Experiment 3 was 2631 ms ($SE = 79$ ms).

Discussion

The attention-switching account of the interaction between syntactic and arithmetic complexity that was observed in Experiments 1 and 2 predicted a similar interaction between syntactic and spatial integration task complexity in Experiment 3. No comparable interaction was observed.

There are at least four possible reasons for why one might not observe an interaction between the language-processing task and the spatial integration task in Experiment 3. First, the spatial integration task might have been too easy, with the consequence that participants were not overly taxed in the condition where the complexity of both tasks was high. A prediction of this hypothesis is that the spatial integration task should be easier to process than the arithmetic-processing task, because the arithmetic-processing task in Experiments 1 and 2 interacted with the language-processing task (either due to a shared pool of working memory resources, or due to the attention-switching costs). Contrary to this prediction, the Norming Study established that the spatial integration task was not easier than the arithmetic task. As discussed above, it took participants longer to perform the spatial integration task than the arithmetic task when the tasks were presented in isolation, and the accuracy on the arithmetic task was at ceiling, with the consequence that the accuracy on the spatial integration task could not be higher. Thus, the lack of an interaction in Experiment 3 was not due to the low complexity of the spatial integration task.

Second, it is possible that the spatial integration task was too difficult, with the consequence that the difficulty of the spatial integration task would swamp the syntactic complexity effect. If this were the case, the following patterns of data would be predicted: either (1) no syntactic complexity effect in the hard spatial integration conditions, or (2) no syntactic complexity effect in both the easy and the hard spatial integration conditions. Contrary to this prediction, our data revealed a main effect of syntactic complexity which was present in both the easy spatial integration task conditions ($Fs > 34$, $ps < .001$) and the hard spatial integration task conditions ($Fs > 13$, $ps < .001$). Thus, the lack of an interaction in Experiment 3 was not due to the high complexity of the spatial integration task.

Third, it is possible that Experiment 3 did not have enough power to detect the interaction between syntactic complexity and spatial integration task complexity. The standard practice in performing post hoc power analyses is to estimate the expected effect size (f) in the experiments at question based on the effect sizes observed in similar experiments in previously conducted research (e.g., Rosenthal & Rosnow, 1991). However, because of the novelty of the experimental paradigm used in the experiments reported here, there were no prior similar studies from which we could estimate the expected effect size for our experiments. We therefore estimated the effect size based on Experiments 1 and 2, where we observed the critical interaction. The f-values for these two experiments—calculated using the partial eta squared values for the interaction effect in each of the experiments—were .307 and .556. To perform the power analysis for the spatial-task experiments (Experiment 3 and Experiment 4 to be presented below), we calculated the mean f-value for Experiments 1 and 2 (.438). The resulting power levels (calculated using G*Power, available at http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/) were as follows: the power in Experiment 3 was .932, and the power in Experiment 4 was .809. These power levels are higher than the power threshold of .80 accepted as a standard in the field (e.g., Cohen, 1988; Rosenthal & Rosnow, 1991). We therefore conclude that our spatial-task experiments (Experiments 3 and 4) had sufficient power to detect an interaction of the size observed in Experiments 1 and 2.

Finally, the lack of an interaction could result from the fact that linguistic and spatial integrations rely on independent pools of working memory resources. Because the first three reasons are not likely to be able to account for the lack of an interaction in Experiment 3, as discussed above, the independence of resource pools for linguistic vs. spatial integration processes is plausibly responsible for the lack of an interaction.

An anonymous reviewer of an earlier version of this paper pointed out an additional concern in comparing the arithmetic task and the spatial integration task. Specifically, it is possible that in the arithmetic task participants might integrate the addends on-line as they go along in the sentence, but in the spatial integration task they might store the pie-slices and add them up at the end of the sentence, thus not performing the integrations on-line, as instructed. If that were the case, the on-line reaction time data from the two tasks (additions vs. the spatial integration task) would not be very meaningful, as the underlying processes involved in the performance of the two tasks would be drastically different. Two sources of evidence suggest that it is unlikely that participants were following this proposed strategy. First, the storing strategy predicts that reaction times should increase from Region 1 to Region 4, peaking at Region 4 where participants would be holding on to three pie-slices from the previous regions and adding the fourth one to the stack. However, this is not the pattern of reaction times we observed in Experiment 3: reaction times are slowest for Regions 2 and 3, but at Region 4 they come down to the reaction time level of Region 1 (see Fig. 6). Second, in the post-experimental questionnaire mentioned above, we coded the subjects' responses for whether they contained any mention of the type of strategy with regard to the time course of performing the task. This constituted 43% of the subjects. Out of these responses, 93% strongly suggested that the task was performed incrementally. Examples of such responses included "made mental hash marks on the circle to keep track of how much space was covered by the blue at each step", "adding past slices as the new slices were added", and "added the pieces together as I went", etc. The remaining 7% of the responses were unclear with regard to the time-course issue. None of the subjects explicitly mentioned performing the task by storing individual pie slices while moving along the sentence and adding up the slices at the end. However, it is difficult to conclusively rule out the possibility that on some trials some participants may have used the storage-based strategy, which could have contributed to the lack of an interaction in the critical region. In summary, based on the pattern of results in reaction times and on the questionnaire responses, we conclude that participants usually performed the spatial integration task incrementally on-line.

Experiment 4

In order to evaluate the generality of the results from Experiment 3, we investigated the relationship between the working memory resources used for on-line language processing and on-line spatial integration processing using a different version of a spatial integration task in a similar dual-task paradigm. In this task, participants were presented with a series of three-by-four grids with some squares filled in in blue. Participants were instructed to imagine combining the squares into a geometrical shape, as shown in Fig. 7.

This spatial integration task is similar to the arithmetic tasks in Experiments 1 and 2 in that an incoming element/incoming elements—a square/squares—must be

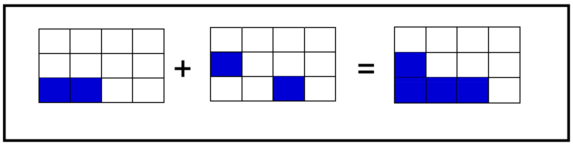

Fig. 7. Spatial integration task in Experiment 4. (For interpretation of the references to colour in the text citation, the reader is referred to the web version of this paper.)

integrated into, or added to, the evolving representation. Similar to the spatial integration task in Experiment 3, the spatial integration task does not rely on verbal working memory resources, and should not therefore interact with the sentence-processing task if the cause for the observed interactions in Experiments 1 and 2 is an overlap in the use of verbal working memory resources.

Methods

Participants

Forty-four participants from MIT and the surrounding community were paid for their participation. All were native speakers of English and were naive as to the purposes of the study. None participated in Experiments 1, 2 or 3.

Design and materials

The experiment had a 2×2 design, crossing syntactic complexity (subject-/object-extracted RCs) with the complexity of the spatial integration task (simple integrations with one square per grid/complex integrations with two squares per grid). The language materials were the same as those used in the other experiments.

The squares for the spatial integration task were randomly generated on-line for each participant in the following way: one square or two adjacent (sharing sides) squares were shown in the first grid; in each subsequent grid, the square(s) that were shown in the previous grid were hidden, and one or two squares were added, such that it/they shared sides with the square(s) in the previous grid.

Procedure

The procedure was identical to that of the other experiments, except for the new spatial integration task. Above each sentence fragment, participants saw a three-by-four grid. The first grid had one or two squares filled in in blue. If two squares were filled in, they were always adjacent. On the next grid, the square(s) that were filled in in the first grid were hidden and one or two other squares were filled in, such that if the square(s) from the first grid were shown, the new square(s) would share sides with them, as shown in Fig. 8. The participants were instructed to imagine constructing a geometrical shape out of the squares.

To assure that the participants performed the task, at the end of each trial a blank grid appeared at the center of the screen. Participants were instructed to click on the squares (using the mouse) that have been highlighted across the four grids. If all the squares were filled in correctly, the word “Right!” flashed briefly on the screen; if all but one square were filled in correctly (including either a false positive or a missing square), the word “Almost” flashed briefly; finally, if the answer was two or more squares off, the words “Not quite” flashed briefly on the screen. As in the other experiments, this task was followed by two comprehension questions about the content of the sentences.

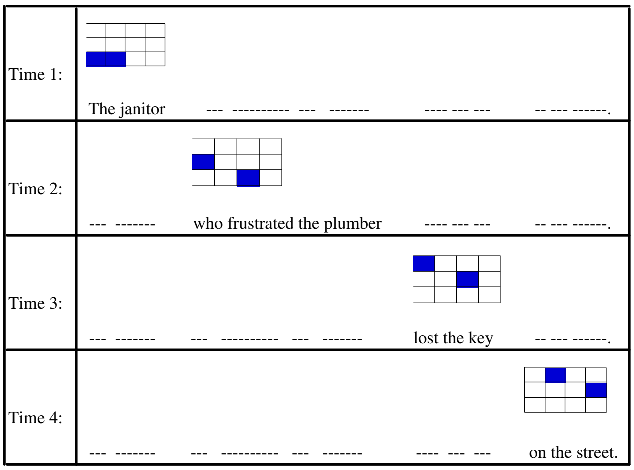

Fig. 8. Sample frame-by-frame presentation of an item in Experiment 4. (For interpretation of the references to colour in the text citation, the reader is referred to the web version of this paper.)

Results

Spatial integration task accuracy

The performance on the spatial integration task was measured in the following way: the number of errors was divided by the total number of squares that should have been selected, where the number of errors were calculated as the maximum of the number of misses and false alarms (i.e., leaving a square out, adding an extra square, or swapping a correct square for an incorrect square would all count as a single error). On average, participants made 5.7% errors. Table 10 presents the mean accuracies (in percent of errors) across the four conditions of Experiment 4. A two-factor ANOVA crossing spatial integration task complexity (easy, hard) and syntactic complexity (easy, hard) revealed a main effect of the spatial integration task complexity ($F1(1,43) = 44.4$; $MSe = 0.0931$; $p < .001$; $F2(1,31) = 97.8$; $MSe = 0.0677$; $p < .001$; min $F'(1,71) = 30.5$; $p < .001$), but no other significant effects (Fs < 2.5).

Comprehension question performance

There were two comprehension questions following each experimental trial. Participants answered the first question correctly 81.9% of the time, and the second question 78.3% of the time. As in the other experiments, we collapsed the results in our analyses. Table 11 presents the mean accuracies across the four conditions of Experiment 4. A two-factor ANOVA crossing spatial integration task complexity (easy, hard) and syntactic complexity (easy, hard) on the responses to the comprehension questions revealed a main effect of the spatial integration task complexity ($F1(1,43) = 25.39$; $MSe = 0.2784$; $p < .001$; $F2(1,31) = 9.71$; $MSe = 0.2025$; $p < .005$; min $F'(1,54 = 7.02$; $p < .02$)), but no other significant effects and no interaction (Fs < 1).

Table 10
Spatial-task accuracies in percent of errors, as a function of spatial task complexity and syntactic complexity (standard errors in parentheses)

Spatial task complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy integrations	2.89 (.71)	3.48 (.86)
Hard integrations	7.78 (1.1)	8.24 (1.2)

Table 11
Comprehension accuracies in percent correct, as a function of spatial task complexity and syntactic complexity (standard errors in parentheses)

Spatial task complexity	Syntactic complexity	
	Subject-extraction (Easy)	Object-extraction (Hard)
Easy integrations	84.2 (2.5)	83.9 (2.1)
Hard integrations	77.1 (2.6)	75.1 (2.3)

Reaction times

As in the other experiments, we analyzed all trials, regardless of how the comprehension questions were answered and how the spatial integration task was performed. Also, as in the previous experiments, reaction time data points that were more than three standard deviations away from the mean residual RT for a position within a condition were excluded from the analyses, affecting 1.5% of the data. Fig. 9 presents the mean residual reaction times per region across the four conditions in Experiment 4.

We conducted a 2×2 ANOVA crossing syntactic complexity and the complexity of the spatial integration task for each of the four regions. The results are presented in Table 12. We first present the analysis of the critical region, Region 2, which included the relative clause ("who frustrated the plumber"/"who the plumber frustrated"), followed by the analyses of the other three regions. At the critical region, the hard-spatial-task conditions were read significantly slower than the easy-spatial-task conditions (326.7 ms vs. −536.8 ms; 95% CI = 160.4 ms), and the syntactically more complex object-extracted RC conditions were read significantly slower than the subject-extracted RC conditions (73.2 ms vs. −283.3 ms; 95% CI = 121.8 ms). Critically, there was no trace of an interaction between syntactic complexity and the complexity of the spatial task. This pattern of results is similar to that in Experiment 3, and it provides additional evidence against the attention-switching explanation of the interaction that was observed in Experiments 1 and 2.

In Region 1, there was an effect of the spatial task complexity, which was not significant in the items analysis. There were no other effects and no interaction. In Region 3, there was a main effect of spatial-task complexity. There was also an unpredicted interaction, such that in the easy-spatial-task conditions, the object-extracted condition is numerically slower, and in the hard-spatial-task conditions, the object-extracted condition is numerically faster than the subject-extracted condition. There is no reason to expect an interaction of this sort here, as the linguistic materials were exactly the same in the subject- and object-extracted conditions. Finally, in Region 4, there was again a main effect of spatial-task complexity, but no other significant effects.

Discussion

The pattern of results in Experiment 4 was very similar to the pattern of results in Experiment 3. Specifically, we found a main effect of linguistic complexity, a

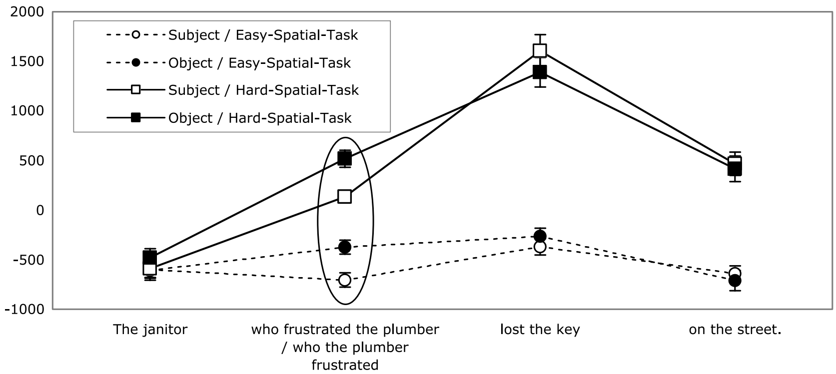

Fig. 9. Reaction times per region in the four conditions of Experiment 4. Error bars indicate standard errors. The critical region is circled.

Table 12
Analysis of variance results for Experiment 4

Source of variance	By participants			By items		min F'	
	df	$F1$ value	*MSe*	*df*	$F2$ value	*df*	min F'
Region 1							
Syntactic complexity	1,43	1.70	69,232	1,31	<1	1,52	<1
Spatial-task complexity	1,43	4.99*	44,118	1,31	2.48	1,59	1.65
Interaction	1,43	2.74	55,747	1,31	1.23	1,57	<1
Region 2 (critical region containing the relative clause)							
Syntactic complexity	1,43	34.86*	160,378	1,31	45.87*	1,74	19.8*
Spatial-task complexity	1,43	117.82*	278,444	1,31	318.85*	1,68	86.0*
Interaction	1,43	<1	105,178	1,31	<1	1,72	<1
Region 3							
Syntactic complexity	1,43	<1	296,470	1,31	<1	1,64	<1
Spatial-task complexity	1,43	97.72*	1481,332	1,31	587.36*	1,56	83.7*
Interaction	1,43	4.25*	266,769	1,31	5.09*	1,74	2.31
Region 4							
Syntactic complexity	1,43	1.26	122,314	1,31	<1	1,57	<1
Spatial-task complexity	1,43	51.32*	1064,008	1,31	584.14*	1,50	47.1*
Interaction	1,43	<1	171,070	1,31	<1	1,53	<1

Note. Significant effects are marked by asterisk.

main effect of spatial integration complexity, but no trace of an interaction. The attention-switching account of the interaction between syntactic and arithmetic complexity that was observed in Experiments 1 and 2 predicted a similar interaction between syntactic and spatial integration task complexity in Experiments 3 and 4. No such interaction was observed in either Experiment 3 or 4. Therefore, we conclude that the interactions observed in Experiments 1 and 2 cannot be accounted for in terms of the attention-switching account.

General discussion

We reported the results of four dual-task experiments which were aimed at investigating the nature of working memory resources in linguistic integrations. The way we approached this question was by crossing syntactic complexity with the complexity of another task, which involved similar integration processes. This secondary task either involved arithmetic integration processes and therefore relied on the use of verbal working memory, or it involved spatial integration processes. Experiments 1 and 2 crossed syntactic complexity and arithmetic complexity. Both of these experiments showed two main effects and a super-additive interaction during the critical region of the linguistic materials, such that in the condition where both syntactic and arithmetic complexity was high the reaction times were longer than would be expected if the two complexity effects were additive. This pattern of results suggests that linguistic and arithmetic integrations rely on overlapping pools of verbal working memory resources.

To account for a potential confound in terms of attention-switching costs in dual-task paradigms, Experiments 3 and 4 crossed syntactic complexity and the complexity of a spatial integration task. The attention-switching account predicts a similar interaction regardless of the nature of the tasks involved, as long as the tasks are matched for complexity. In contrast, the hypothesis whereby linguistic and arithmetic integrations rely on overlapping pools of verbal working memory resources predicts no interaction in cases when one of the tasks does not rely on verbal working memory resources. Both Experiment 3 and Experiment 4 revealed two main effects, but no suggestion of an interaction comparable to the interactions observed in Experiments 1 and 2. These results therefore provide evidence against the attention-switching account.

We discussed three alternative hypotheses which could account for the lack of an interaction in the spatial task experiments: (1) the spatial tasks may have been too easy, (2) the spatial tasks may have been too hard, and (3) the spatial task experiments may not have had enough power to detect an interaction similar to the one observed in Experiments 1 and 2. We presented arguments against each of these hypotheses. By comparing the arithmetic task used in Experiment 1 and the spatial integration task used in Experiment 3, we first established that the spatial integration task was not too easy (1) by showing that in the Norming Study—where each of these tasks was performed in isolation by an independent group of participants—the spatial integration task took longer to perform than the arithmetic task and the accuracies in the spatial integration task were plausibly lower; and (2) by showing that in Experiments 1 and 3, the ranges of raw reaction times were very similar across all regions and at the critical region. Second, we established that the spatial integration task was not too difficult by showing that in Experiment 3, a significant main effect of syntactic complexity was observed, indicating that the spatial integration task was not swamping the syntactic complexity effect. Finally, we established that it was likely that Experiments 3 and 4 had sufficient power to detect an interaction of the size observed in Experiments 1 and 2: the power analysis revealed that the power in both Experiments 3 and 4 was >.80. We therefore argued that in the spatial-task experiments, the lack of an interaction similar to the interaction observed in the arithmetic task experiments was plausibly due to the fact that whereas linguistic and arithmetic integration processes rely on overlapping pools of verbal working memory resource, linguistic and spatial integration processes do not, at least not to the same degree.

In the reaction time analyses for the four experiments above, we focused on the critical region (Region 2 in all the experiments) where linguistic complexity was manipulated. It is worth noting, however, that the overall data patterns differ, to some extent, across the four experiments. Whereas there is a main effect of secondary task complexity on Regions 2–4 in all four experiments, the reaction times peak at different regions: at Region 2 in Experiments 1 and 3, and at Region 3 in Experiments 2 and 4. Importantly though, these differences in the peak point of reaction times do not correlate with the type of the secondary task: specifically, the secondary task in Experiment 1 is the *arithmetic* addition task, while in Experiment 3 it is the *spatial* pie task; similarly, the secondary task in Experiment 2 is the *arithmetic* addition–subtraction task, while in Experiment 4 it is the *spatial* grid task. Because (1) we attempted to generalize over the two arithmetic tasks and the two spatial tasks, and distinguish between the arithmetic and spatial tasks, and (2) the critical interaction was observed during Region 2 in Experiments 1 and 2 and not in any of the other regions in any of the four experiments, the differences in the peak reaction times between Experiments 1 and 3 on one hand and Experiments 2 and 4 on the other hand do not seem relevant to the interpretation of the critical contrast between the presence of a super-additive interaction in Experiments 1 and 2 (with the arithmetic secondary tasks) and the absence of such an interaction in Experiments 3 and 4 (with the spatial secondary tasks). We hypothesize that the differences in the peak reaction times across experiments may be resulting from the differences in the difficulty of the secondary tasks across the experiments: the secondary tasks in Experiments 2 (addition/subtraction) and 4 (the grid spatial task) are plausibly more difficult than the secondary tasks in Experiments 1 and 3.

In our experimental logic we relied on the assumption that verbal and visuo-spatial working memory resource pools are independent, based on the earlier studies. The evidence for the independence of these two working memory resource pools comes from several kinds of studies: (1) dual-task experiments showing selective interference effects, such that a verbal memory task interferes to a larger extent with another verbal memory task, compared to a spatial memory task, and vice versa (e.g., Baddeley, 1986; Logie, 1986, 1995); (2) individual-differences studies showing that the correlations in people's performance are higher within domains (verbal or visuo-spatial), than across domains (e.g., Shah & Miyake, 1996); (3) neuropsychological case studies of patients who are selectively impaired on verbal memory tasks or spatial memory tasks (Hanley et al., 1991; Vallar & Shallice, 1990); and (4) neuroimaging studies suggesting that different neural substrates underlie verbal memory tasks and spatial memory tasks (Jonides et al., 1993; Paulesu et al., 1993). All these different lines of evidence converge in their conclusions that there exist separate resource pools for verbal vs. visuo-spatial memory. It is worth noting, however, that in some of the previous studies the verbal and the visuo-spatial memory tasks were quite different in terms of the cognitive processes they involve, i.e., the tasks differed in more respects than the use of verbal vs. visuo-spatial resources. For example, a standard manipulation in dual-task experiments comparing the degree of interference produced by tasks from different domains has involved tapping the four corners of a square with a finger continuously for the spatial distractor task, and repeatedly pronouncing a word for the verbal distractor task. Even though we did not intend to test the hypothesis that verbal and visuo-spatial working memory resource pools are independent (we assumed this to be the case, based on the earlier evidence), our results can be taken as additional strong evidence for the independence of these two resource pools. Specifically, in our experiments the arithmetic tasks (which rely on verbal working memory resources) and the spatial integration tasks (which rely on spatial working memory resources) were qualitatively very similar in terms of the cognitive processes they involved (combining representations into more complex representations over time), and yet they showed differential interference with respect to the language-processing task, which relies on verbal working memory resources.

As discussed in the Introduction, there exists behavioral and ERP evidence for two different types of working memory costs in on-line language processing: working memory resources for processing incomplete syntactic dependencies (Chen et al., 2005; Chomsky & Miller, 1963; Gibson, 1998, 2000; Kluender & Kutas, 1993; Lewis, 1996; Wanner & Maratsos, 1978), and working memory resources for integrating words to earlier positions in the sentence (Gibson, 1998, 2000; Gordon et al., 2001; Grodner & Gibson, 2005; Kaan et al., 2002; Phillips et al., 2005; Warren & Gibson, 2002). We argued that it might be necessary to take this evidence into consideration when investigating the extent of domain-specificity of working memory resources for on-line language processing. Specifically, we suggested that the two pools of working memory resources used in on-line language processing—the one involved in keeping track of incomplete syntactic dependencies and the one involved in integrating structural elements over long distances—may differ in the extent of their domain-specificity and in the extent of their overlap with other working memory systems. In this paper, we focused on investigating the nature of working memory resources in linguistic integrations by examining the relationship between linguistic integrations and similar integration processes which either involve or do not involve verbal working memory resources. We provided evidence for an overlap in resource pools used for linguistic and arithmetic integration processes. This suggests that future investigations aimed at understanding the nature of working memory resources in language processing may in fact benefit from examining the two different resource pools used in language processing independently.

The results reported here may be used to suggest that the verbal working memory resource pool is divided along the lines of the qualitative nature of the cognitive processes involved, rather than along the domains to which the tasks belong. For example, Caplan and Waters (1999) argued that the verbal working memory resource pool is divided into resources used for on-line language processing and resources used for non-linguistic verbally mediated tasks. However, it is possible that different pools of verbal working memory resources are used (1) for tasks which involve storing verbal representations in memory over time, and (2) for tasks which involve combining verbal representations into more complex representations. The behavioral and ERP evidence from the language-processing literature discussed above is consistent with this line of reasoning, such that even within the resource pool for on-line language processing there appear to exist two independent sub-pools of resources—for keeping track of incomplete syntactic dependencies and for integrating structural elements with one another. Therefore, in conjunction with the results reported here, it is plausible that the resource pool that any given verbal task may rely on is determined by the nature of the processes involved in the task.

Acknowledgments

This research was supported by NIH Grant 5R01DC05237-05 to the second author and NIH NRSA Grant 1-F32-MH65105-02 to the third author. We would like to thank the following people for helpful discussions of this work: Asaf Bachrach, David Badre, Mara Breen, David Caplan, Daniel Casasanto, Sara Cordes, Timmy Desmet, Florian Jaeger, Yuhong Jiang, Nancy Kanwisher, Josh McDermott, Maryellen MacDonald, Neal Pearlmutter, Molly Potter, Adam Szczegielniak, Paul Smolensky, Gloria Waters, Ken Wexler, Jon Winawer, Florian Wolf, and the audiences at the CUNY conference in 2004 and 2005, and at the Cognitive Science Conference 2004.

Appendix A. Linguistic materials

One of the four subject-/object-extracted RC versions is shown below for each of the 32 items. The other three versions can be generated as exemplified in (1) below.

(1)

a. Subject-extracted, version 1:
The janitor who frustrated the plumber lost the key on the street.

b. Subject-extracted, version 2:
The plumber who frustrated the janitor lost the key on the street.
c. Object-extracted, version 1:
The janitor who the plumber frustrated lost the key on the street.
d. Object-extracted, version 2:
The plumber who the janitor frustrated lost the key on the street.

(2) The hairdresser who hired the beautician transformed the salon for the better.
(3) The lecturer who provoked the dean left the university in the summer.
(4) The trumpeter who loved the drummer formed the band two years ago.
(5) The intern who distrusted the boss disregarded the messages on her voicemail.
(6) The roommate who annoyed the landlord slammed the door of the apartment.
(7) The player who avoided the coach entered the room at the gym.
(8) The mayor who called the advisor requested an update on the project.
(9) The librarian who angered the teacher misplaced the book from the depository.
(10) The pharmacist who helped the assistant placed the order for the drug.
(11) The waitress who hugged the bartender dropped the tray on the floor.
(12) The client who contacted the retailer offered a deal of the century.
(13) The celebrity who admired the athlete won the award at the ceremony.
(14) The detective who recognized the spy crossed the street at the light.
(15) The journalist who complimented the editor revised the article for the newspaper.
(16) The employee who praised the executive finished the project right on time.
(17) The legislator who visited the senator falsified the documents for the trip.
(18) The soldier who shot the enemy received a medal for the battle.
(19) The officer who described the murderer told a lie about the past.
(20) The reporter who followed the cameraman damaged the equipment during the trip.
(21) The understudy who telephoned the agent shared the news about the suicide.
(22) The consultant who confronted the programmer broke the computer in a rage.
(23) The supervisor who deceived the owner kept the money in the end.
(24) The entrepreneur who hated the stockbroker sold the shares after the merger.
(25) The mole who revealed the defector rejected the offer on the spot.
(26) The singer who blamed the organizer cancelled the concert in Los Angeles.
(27) The acrobat who mocked the clown performed the trick at the show.
(28) The customer who upset the seller forgot the receipt on the counter.
(29) The partner who introduced the businessman presented the report at the meeting.
(30) The messenger who summoned the knight read the letter from the king.
(31) The linguist who ridiculed the historian proposed the hypothesis for the problem.
(32) The biker who ignored the driver made the turn at the crossing.

References

Ashcraft, M. H. (1992). Cognitive arithmetic: a review of data and theory. *Cognition, 44*, 75–106.

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: a review and summary of new directions. *Mathematical Cognition, 1*, 3–34.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

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

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*(2), 299–315.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Brain and Behavioral Sciences, 22*, 77–126.

Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language, 52*(1), 144–169.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.). *Handbook of mathematical psychology* (Vol. 2, pp. 269–321). New York: Wiley.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

Donders, F. C. (1868–1869). Over de snelheid van psychische processen. Onderzoekingen gedann in het Physiologish Laboratorium der Utrechtsche Hoogeshool. Tweede Reeks, II, 92–120. In W. G. Koster (Ed. and Transl.), *Attention and performance II*. Amsterdam: North-Holland, 1969. (Reprinted from *Acta Psychologica, 30*, 412–431 (1969))

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, 54.

Felser, C., Clahsen, H., & Munte, T. F. (2003). Storage and Integration in the processing of filler-gap dependencies: an ERP study of topicalization and wh-movement in German. *Brain and Language*, 87.

Ferreira, F., & Clifton, C. Jr., (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348–368.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. *Journal of Memory and Language, 47*(2), 250–272.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 203–218.

Gibson, E. (1991). *A computational theory of human linguistic processing: memory limitations and processing breakdown*. Unpublished doctoral dissertation. Pittsburgh, PA: Carnegie Mellon University.

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

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain*. Cambridge, MA: MIT Press.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1411–1423.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13*, 425–430.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*, 97–114.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science, 29*(2), 261–290.

Hanley, J. R., Young, A., & Pearson, N. A. (1991). Impairment of the visuo-spatial scratchpad. *Quarterly Journal of Experimental Psychology, 43A*, 101–125.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature, 363*, 623–625.

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

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processing in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228–238.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2002). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159–201.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language, 30*, 580–602.

King, J., & Kutas, M. (1995). Who did what to when?: Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7*(3), 376–395.

Kluender, R., & Kutas, M. (1993). Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*(2), 196–214.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research, 29*(6), 627–645.

Lewis, R. (1996). A theory of grammatical but unacceptable embeddings. *Journal of Psycholinguistic Research, 25*, 93–116.

Loftus, G. (1978). On interpretation of interactions. *Memory & Cognition, 6*(3), 312–319.

Logie, R. H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 229–247.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, UK: Erlbaum.

Masson, M. E., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 202–220.

McClelland, J. L. (1979). On the time relations of mental processes: an examination of systems of processing in cascade. *Psychological Review, 86*(4), 287–330.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language, 48*, 67–91.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*, 342–345.

Phillips, C., Kazanina, N., & Abada, S. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research, 22*, 407–428.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York, NY: McGraw-Hill Series in Psychology.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual differences approach. *Journal of Experimental Psychology: General, 125*(1), 4–27.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.

Siegler, R. S. (1987). Strategy choices in subtraction. In J. A. Slobada & D. Rogers (Eds.), *Cognitive processes in mathematics*. Oxford: Clarendon Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North-Holland (Reprinted from *Acta Psychologica, 30*, 276–315 (1969).

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33*, 285–318.

Vallar, G., & Shallice, T. (Eds.). (1990). *Neuropsychological impairments of short-term memory*. New York: Cambridge University Press.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: a cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49*, 285–316.

Vasishth, S. (2003). *Working memory in sentence comprehension*. New York: Garland Press.

Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, G. Miller, & J. Bresnan (Eds.), *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition, 85*, 79–112.

Waters, G., & Caplan, D. (1999). *The effect of a digit-load on syntactic processing in subjects with high and low working memory capacity*. Unpublished manuscript.

Waters, G., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.

Waters, G., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology, 12*, 1–30.