On-line syntactic processing under concurrent memory load

GLORIA WATERS

Boston University, Boston, Massachusetts

DAVID CAPLAN

Massachusetts General Hospital, Boston, Massachusetts

and

SASHA YAMPOLSKY

Boston University, Boston, Massachusetts

Thirty-six university students were tested in a plausibility judgment task using a self-paced listening paradigm under no-interference and two-digit load conditions. Listening times were longer at syntactically more complex portions of syntactically more complex sentences, and greater loads led to increased listening times. However, listening times at syntactically more complex positions in syntactically more complex sentences did not increase more than listening times at comparable positions in syntactically simple sentences under digit load conditions. The results indicate that a concurrent memory load does not reduce the availability of working memory resources used for on-line syntactic processing and, thus, provide evidence that the working memory system used for assigning syntactic structure is separate from that measured by standard working memory tasks.

Assigning syntactic relationships to words in sentences involves performing computational operations on items retained for short periods of time in memory, as can be seen in the need to assign *the boy*, not *the girl*, to the grammatical role of the subject of *fell* in *The boy who chased the girl fell*. The nature of the working memory system involved in this processing is controversial. Some researchers argue that there is a single working memory system that is used for all verbal tasks (Just & Carpenter, 1992; King & Just, 1991). Others claim that determining the meaning of a sentence on the basis of its syntactic structure involves a specialized working memory system that is separate from that used in other verbal tasks (Caplan & Waters, 1999).

One approach to investigating the nature of the working memory system involved in syntactic processing has been to examine the effects of a concurrent memory load on syntactic processing. The rationale for this approach is that, if there is a single verbal working memory system, engaging that system with a concurrent task should make syntactic processing less efficient. Since a concurrent load would be expected to interfere with many aspects of sentence processing, simply finding that participants made more errors or had longer reaction times (RTs) in a sentence comprehension task under concurrent load conditions would not specifically indicate that the load inter-

fered with syntactic processing. Evidence for a specific effect of load on syntactic processing would come from the finding that processing syntactically more complex sentences suffers disproportionately, as compared with the processing of syntactically simple sentences under load conditions (see Caplan & Waters, 1999, for more discussion).

Studies in which concurrent memory loads have been used have generally measured sentence processing off line, where responses made by participants occur after the entire sentence has been presented. Overall, end-of-sentence measures have not found strong evidence that a concurrent memory load interferes with syntactic processing, but the studies are not definitive, for several reasons. Baddeley and Hitch (1974) reported that maintaining a six-digit load interfered with making judgments as to whether a sentence about sequences of letters matched a visual depiction of those sequences more for passive than for active sentences. The interaction between digit load and syntactic complexity of the sentence suggests that a memory load interferes with syntactic processing, but many features of the task complicated the study—half of the trials presented familiar alphabetical sequences, whereas half contained a negation—and it is not known whether the active/passive difference was confounded in any way with these features of the stimuli. Waters, Caplan, and Hildebrandt (1987) reported an interaction between syntactic complexity and digit load in a plausibility judgment task, but the interaction was significant only in the analysis of results by subjects, making the generalizability of the result questionable. Dick et al. (2001) found that a digit load did

Correspondence concerning this article should be addressed to G. Waters, Department of Health Sciences, Boston University, 635 Commonwealth Ave., Boston, MA 02215 (e-mail: gwaters@bu.edu).

not have an effect on the accuracy of participants' abilities to report the actor in a sentence, but these results are also hard to interpret, since the participants' performance was at ceiling in all conditions. Waters, Caplan, and Rochon (1995) found that a concurrent digit load did not affect comprehension of syntactically more complex sentence types in either patients with Alzheimer's disease or control participants.

More sensitive tests of syntactic processing involve on-line measures of the moment-by-moment construction of syntactic form as the words in a sentence are presented. Techniques such as measurement of eye fixation duration, self-paced word-by-word and phrase-by-phrase reading time, and event-related potentials (ERPs) have shown that on-line processing becomes more difficult at specific portions of sentences. These portions include words that resolve ambiguities (Ferreira & Clifton, 1986; Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996; Osterhout & Holcomb, 1992) and words that require integration with many other elements in a sentence, especially when these elements have occurred much earlier in a sentence or are separated from the current word by certain types of phrases (Gibson, 1998). Very few studies have measured on-line syntactic processing under concurrent load conditions. Vos, Gunter, Schriefers, and Friederici (2001) reported ERP effects of an increased load (concurrent detection of one of three words vs. detection of a single word presented before a sentence) on processing syntactically more complex sentences at the point at which processing load increased. However, the ERP effect of syntactic complexity occurred in a late wave that the authors associated with re-analysis, not with initial assignment, of syntactic structure. Therefore this effect is hard to interpret as evidence that a concurrent load competes with on-line syntactic processing for processing resources. Vos et al.'s (2001) paradigm did not allow on-line behavioral measures to be recorded, since words were presented at fixed intervals to allow ERPs to be measured. To our knowledge, the only study that has measured on-line behavioral responses to sentences that differed in syntactic complexity under varying load conditions is that by King and Just (1991), who measured self-paced reading time for phrases in sentences that differed in syntactic complexity. However, these authors focused on the reading times in participants with different working memory spans and did not report the results for a no-load condition, as compared with the load conditions, on phrase-by-phrase reading times in the different sentence types.

Exploring the effect of load on on-line syntactic processing offers a potentially sensitive approach to the question of the nature of the working memory system involved in syntactic processing. The present article reports data from such a study.

METHOD

Participants

Thirty-six university students, 18–30 years of age, recruited through ads posted in the university, participated in the study. The partici-

pants were required to have English as their native language, no history of neurological or psychiatric disease, and normal hearing.

Procedures and Materials

On-line measure of sentence-processing efficiency. The auditory moving windows (AMW) paradigm (Ferreira et al., 1996) with an end-of-sentence plausibility judgment task (to ensure comprehension) was used as the on-line sentence-processing task. The materials were constructed so that the plausibility judgment task could not be performed correctly unless the participants assigned the syntactic structure of the sentences and processed them for meaning (see below). The participants heard sentences that were segmented into phrases, and they pressed a button on a box interfaced with the computer to hear each successive phrase. RTs for each phrase were recorded. Half of the sentences were semantically plausible and half were implausible. The participants heard a tone at the offset of the sentence-final word to indicate the termination of the sentence. Plausibility judgments were made upon hearing the tone by pressing a *yes* or a *no* button, and judgment RT and judgment accuracy were recorded. All the participants participated in three conditions (no load, three-digit load, and five-digit load) over the course of three sessions, which were counterbalanced across participants.

Digit load conditions. In the digit load conditions, the participants were required to remember numbers while performing the AMW task. The participants first saw the word "ready" on the screen and pressed a button on the button box to trigger presentation of the digits on the screen. Length of digit presentation was determined via pilot testing and was 900 msec for the three-digit load and 1,500 msec for the five-digit load. Following presentation of the digits, the self-paced task began. After making the plausibility judgment, the participants heard a short beep, and a set of question marks ("??") appeared on the screen. This indicated to the participant that he or she should repeat the list of numbers that were seen at the beginning of the sentence. The participants were instructed to produce as many numbers as possible in the order that they had seen them. The experimenter recorded the participant's digit recall response on a score sheet. The word "ready" appeared on the screen again, and the sequence of events was repeated. In an effort to reduce individual differences in tradeoffs in attending to the sentence and the recall portions of the task, the participants were instructed that the first priority should be to remember the digits correctly and then to perform the sentence task as quickly as possible.

Stimuli. The stimuli consisted of 26 semantically plausible and 26 semantically implausible sentences of each of the five sentence types shown in Table 1. Sentences were developed in pairs (cleft-subject [CS] and cleft-object [CO], object-subject [OS] and subject-object [SO], and subject-subject [SS] and SO), with the same words in a different word order. The memory storage and computational requirements for processing the second, syntactically more complex, object-relativized member of each pair are greater than those for the corresponding subject-relativized sentences. At the embedded verb of CO and SO sentences, the parser assigns two thematic roles—one to the agent and one to the theme of the verb—whereas at the embedded verb of CS, OS, and SS sentences, the parser assigns only the thematic role of agent. The larger number of *integration* operations at the embedded verb in CO and SO sentences increases the processing resource demands at that point in these sentences (Caplan, Hildebrandt, & Waters, 1994; Gibson, 1998). This effect can spill over to the main verb of SO sentences (King & Just, 1991). Thus, the main data of interest in this study consisted of listening times for the verb in plausible CO sentences, as compared with CS sentences, and for both verbs in plausible SO sentences, as compared with OS and SS sentences. Reading times have been found to be longer for words in sentence-final position (Balogh, Zurif, Prather, Swinney, & Finkel, 1998). To allow for a comparison of listening times for the verb of CO sentences and the verb of CS sentences that was not affected by sentence-final position, a short phrase was added to each of these sentence types. Additional comparisons of interest that con-

Table 1
Examples of Sentence Stimuli

Cleft-Subject (Simple)	Intro	NP1	pro	V	NP2	AdjP
	<i>plausible</i> : /It was/ /the movie/ /that/ /terrified/ /the child/ /because it showed a monster/					
	<i>implausible</i> : /It was/ /the man/ /that/ /delighted/ /the camera in the film/					
Cleft-Object (Complex)	Intro	NP1	pro	NP2	V	AdjP
	<i>plausible</i> : /It was/ /the child /that/ /the movie/ /terrified/ because it showed a monster/					
	<i>implausible</i> : /It was/ /the camera/ /that/ /the man/ /delighted/ /in the film/					
Object-Subject (Simple)	NP1	V1	NP2	pro	V2	NP3
	<i>plausible</i> : /The scout/ /warmed/ /the cabin/ /that/ /contained/ /the firewood/					
	<i>implausible</i> : /The criminal/ /cursed/ /the judge/ /that/ /astonished/ /the verdict/					
Subject-Subject (Simple)	NP1	pro	V1	NP2	V2	NP3
	<i>plausible</i> : /The cabin/ /that /warmed/ /the scout/ /contained/ /the firewood/					
	<i>implausible</i> : /The judge/ /who/ /cursed/ /the criminal/ /astonished /the verdict/					
Subject-Object (Complex)	NP1	pro	NP2	V1	V2	NP3
	<i>plausible</i> : /The cabin/ /that /the scout/ /warmed/ /contained/ /the firewood/					
	<i>implausible</i> : /The judge/ /who/ /the criminal/ /cursed/ /astonished/ /the verdict/					

trol for sentence position are between listening times for the first verb (the verb in the relative clause) of the object-relativized sentences (CO and SO) and the second noun phrase (NP2) of the subject-relativized sentences (CS, SS, and OS), which occur in similar, clause-final positions.

The stimuli were designed so that plausibility judgments did not require judgments based on detailed semantic knowledge. Sentences had verbs that required either animate objects or animate subjects (e.g., *It was the girl that the food nourished*). Implausible sentences violated these constraints. Detection of an anomaly could thus be based on fairly accessible, general semantic features and did not require extensive searches through semantic memory for item-specific information. The animacy of the nouns not affected by animacy constraints of the verb was counterbalanced so that there was no order of animate and inanimate noun phrases that could be used to correctly determine sentence plausibility in any of the four sentence types. The point at which a judgment about implausibility could be made occurred in the first clause in half of the OS, SS, and SO sentences and in the second clause for the other half, to ensure that the participants had to pace their way through the entire sentence before making a decision regarding plausibility.

The sentences were recorded by a male speaker and were digitized at a sampling rate of 40 kHz with a 16-bit quantization by SoundEdit software (Dunn, 1994). The waveform files were then edited (using SoundEdit) to place a marker (or *tag*) at each phrase boundary. Tags were placed in areas of the waveform in such a way that intelligibility was maximized, as determined by perceptual analysis. In addition, visual and perceptual inspections were used to place tags in areas of low signal amplitude whenever possible, in order to make smooth phrase-to-phrase transitions.

The experiment was presented in PsychoScope (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macintosh computer. The wavefiles were entered into PsychoScope and were played out over headphones (after digital-to-analog conversion). PsychoScope recorded (1) the time to play out and process each phrase, (2) the time from the end of the sentence to the participant's buttonpress for plausibility judgments, and (3) the plausibility judgment response.

RESULTS

On-Line Measure of Sentence-Processing Efficiency

The RT for each phrase on the AMW task was the dependent measure of on-line sentence-processing effi-

ciency in all conditions. Each segment's duration was subtracted from the RT to generate *listening times* that were independent of the duration of each segment. Analyses of variance (ANOVAs) for the effects of schedule did not reveal a main effect or interaction with other variables, and this factor was therefore removed from the analyses. Since listening times may have been affected by the frequency of the lexical items in the sentences, the lexical frequencies of each word (Kučera & Francis, 1970) in the positions of theoretical interest were compared in *t* tests and were found not to differ.

The single-resource theory claims that an external load and the load related to complexity of syntactic processing within the sentence utilize the same resource pool, and the separate sentence interpretation resource theory claims that they utilize different pools. Rephrased in terms of expected results, the single-resource theory predicts an interaction between sentence type, phrase, and load, such that listening times at the most demanding point of the more complex sentences will show a greater increase under conditions of external load than will listening times at the comparable point of the less complex sentences. In contrast, the separate sentence interpretation resource theory predicts that there will be an interaction of sentence type and phrase, such that listening times will be longer at the most demanding point of the more complex sentences than at the comparable point of the less complex sentences, but that there will be no further interaction of sentence type and phrase with load; that is, this difference will not increase under conditions of load.

Listening times for plausible sentences for which participants made correct plausibility judgments were analyzed with ANOVAs, using both subject (F_1) and item (F_2) means as units. CS and CO sentences were analyzed in 3 (load: no interference, three digits, or five digits) $\times$ 2 (sentence type) $\times$ 5 (phrase: introduction, first noun phrase [NP1], pronoun, verb, or NP2) analyses. SO and SS, and SO and OS sentences were analyzed with 3 (load: no interference, three digits, or five digits) $\times$ 2 (sentence type)

× 5 (phrase: NP1, pronoun, first verb, NP2, or second verb) analyses.

Figure 1 shows the mean listening time for the three load conditions at each phrase for the three sentence pairs.

For the CS/CO sentences, there were main effects of load [$F_1(2,70) = 26.74$, $MS_e = 47,138.65$, $p < .0001$; $F_2(2,500) = 211.62$, $MS_e = 3,907.5$, $p < .0001$] and phrase [$F_1(4,140) = 98.57$, $MS_e = 15,422.46$, $p < .0001$; $F_2(4,250) = 101.86$, $MS_e = 10,735.2$, $p < .0001$]. The main effect of sentence type was significant in the subject analyses, but not in the item analyses [$F_1(1,35) = 20.55$, $MS_e = 2,734.00$, $p < .0001$; $F_2(1,250) = 2.89$, $MS_e = 10,735.2$, n.s.]. There were significant interactions of sentence type and phrase [$F_1(4,140) = 30.48$, $MS_e = 4,280.92$, $p < .0001$; $F_2(4,250) = 8.84$, $MS_e = 10,735.2$, $p < .0001$] and of load

and phrase [$F_1(8,280) = 7.20$, $MS_e = 4,147.13$, $p < .0001$; $F_2(8,500) = 4.77$, $p < .0001$]. Analysis of the sentence type × phrase interaction with Tukey post hoc tests showed that listening times were longer for the verb in CO sentences than for either the verb or the noun of the relative clause (NP2) in CS sentences. Tukey post hoc tests showed that listening times in the five-digit load condition were longer than those in the three-digit load condition and that those in the three- and five-digit load conditions were longer than those in the no-load condition. The interaction of load and phrase was due to differences in the effects of load on different phrases that were of no theoretical relevance. The load × sentence type interaction and the critical load × sentence type × phrase interaction were not significant in either the subject or the item analyses.

A. Cleft-subject versus cleft-object sentences

B. Object-subject versus subject-object sentences

C. Subject-subject versus subject-object sentences

Figure 1. Mean listening times across load conditions. (A) Cleft-subject and cleft-object sentences. (B) Object-subject and subject-object sentences. (C) Subject-subject and subject-object sentences. Intro, introduction; Pro, pronoun; NP1, first noun phrase; V, verb; NP2, second noun phrase.

For the OS/SO sentences, there were significant main effects of load [$F_1(2,70) = 22.06$, $MS_e = 51,767.55$, $p < .0001$; $F_2(2,500) = 120.5$, $MS_e = 4,755$, $p < .0001$] and phrase [$F_1(4,140) = 79.80$, $MS_e = 15,913.18$, $p < .0001$; $F_2(2,250) = 97.51$, $MS_e = 9,337.2$, $p < .0001$]. The main effect of sentence type was significant in the subject analysis, but not in the item analysis [$F_1(1,35) = 14.14$, $MS_e = 3,341.11$, $p < .001$; $F_2(1,250) = 2.05$, $MS_e = 9,337.2$, n.s.]. There were significant interactions of sentence type and phrase [$F_1(4,140) = 54.15$, $MS_e = 2,612.28$, $p < .0001$; $F_2(4,250) = 11.13$, $MS_e = 9,337.2$, $p < .0001$] and load and phrase [$F_1(8,280) = 4.36$, $MS_e = 4,737.73$, $p < .0001$; $F_2(8,500) = 2.7$, $MS_e = 4,755.4$, $p < .01$]. Analysis of the sentence type $\times$ phrase interaction showed that listening times were longer for the first verb (V1) in SO sentences than for either the verb or the noun of the relative clause (NP2) in SS sentences. Tukey post hoc tests showed that listening times in the five-digit load condition were longer than those in the three-digit load condition and that those in the three- and five-digit load conditions were longer than those in the no-load condition. The interaction of load and phrase was due to differences in the effects of load on different phrases that were of no theoretical relevance. The critical load $\times$ sentence type $\times$ phrase interaction was significant in the subject analysis, but not in the item analysis [$F_1(8,280) = 4.88$, $MS_e = 1,760.84$, $p < .0001$; $F_2(8,500) = 1.45$, $MS_e = 4,755$, n.s.].

For the SS/SO sentences, there were significant main effects of load [$F_1(2,70) = 16.41$, $MS_e = 48,975.35$, $p < .0001$; $F_2(2,500) = 116.15$, $MS_e = 4,424.8$, $p < .0001$], phrase [$F_1(4,140) = 84.67$, $MS_e = 15,697.78$, $p < .0001$; $F_2(2,250) = 106.4$, $MS_e = 8,903.9$, $p < .0001$], and sentence type [$F_1(1,35) = 26.58$, $MS_e = 4,492.55$, $p < .0001$; $F_2(1,250) = 11.84$, $MS_e = 8,903.9$, $p < .001$]. There were significant interactions of sentence type and phrase [$F_1(4,140) = 26.33$, $MS_e = 4,103.65$, $p < .0001$; $F_2(4,250) = 8.69$, $MS_e = 8,903.9$, $p < .0001$] and of load and phrase [$F_1(8,280) = 6.03$, $MS_e = 5,403.93$, $p < .0001$; $F_2(8,500) = 5.09$, $MS_e = 4,424.8$, $p < .0001$]. Analysis of the sentence type $\times$ phrase interaction with Tukey post hoc tests showed that listening times were longer for the verb in SO sentences than for either the verb or the noun of the relative clause (NP2) in SS sentences. Tukey post hoc tests showed that listening times in the five-digit load condition were longer than those in the three-digit load condition and that those in the three- and five-digit load conditions were longer than those in the no-load condition. The interaction of load and phrase was due to differences in the effects of load on different phrases that were of no theoretical relevance. The critical load $\times$ sentence type and load $\times$ sentence type $\times$ phrase interactions were not significant in either the subject or the item analyses.

In summary, the results for all three pairs of sentences (CS vs. CO, OS vs. SO, and SS vs. SO) showed that there was the expected increase in listening time at the verb of the more complex object-relativized sentences, as compared with the simpler subject-relativized sentences. However, as can be seen in Figure 2, the increase in the lis-

Figure 2. Mean increase in listening time for the capacity-demanding portion (the verb) of object- versus subject-relativized sentences in the three load conditions.

tening time difference at the verb did not systematically increase with the size of the concurrent load.

Plausibility Judgments

Figure 3 shows the mean A' scores for the three sentence-pair comparisons (CS/CO, OS/SO, and SS/SO) in the three load conditions. The data were analyzed with 3 (load) $\times$ 2 (sentence type) ANOVAs. For CS and CO sentences, there was a main effect of sentence type [$F_1(1,35) = 32.64$, $MS_e = 0.001$, $p < .0001$]. A' s were higher for the syntactically simpler CS sentences than for the more complex CO sentences. For OS/SO and SS/SO sentences, there were no significant main effects of sentence type. There was no main effect of load and no significant load $\times$ sentence type interaction for any comparison.

Figure 3 also shows plausibility judgment RT data for plausible sentences correctly judged to be plausible and for which digit recall was correct. The data were analyzed with 3 (load) $\times$ 2 (sentence type) ANOVAs with subject and item means as units. For CS and CO sentences, there was a main effect of sentence type [$F_1(1,35) = 5.32$, $MS_e = 91,132.45$, $p < .05$; $F_2(1,50) = 7.36$, $MS_e = 36,841.3$, $p < .01$]. The main effect of load was not significant [$F_1(2,70) = 2.30$, $MS_e = 29,575.04$, n.s.; $F_2(2,100) = 0.87$, $MS_e = 17,459.9$, n.s.]. There was a significant load $\times$ sentence type interaction [$F_1(2,70) = 7.97$, $MS_e = 19,454.83$, $p < .005$; $F_2(2,100) = 6.31$, $MS_e = 17,459.9$, $p < .001$]. A post hoc analysis of the effect of sentence type showed that RTs were faster for CS sentences than for CO sentences. Analysis of the interaction of load and sentence type showed that RTs were shorter for CS sentences than for CO sentences in the no-load condition only and that RTs for the CS sentences were shorter in the no-load con-

A. Plausibility judgment A' s

B. Plausibility judgment (end of sentence) reaction times

C. Percentages of digit recall errors

Figure 3. (A) Mean A' scores for plausibility judgments. (B) Mean plausibility judgment reaction times (RTs). (C) Mean percentages of digit recall errors. CS, cleft-subject; CO, cleft-object; OS, object-subject; SO, subject-object; SS, subject-subject.

dition than in the other two conditions. There were no significant differences in RTs between the two sentence types in either load condition.

For OS/SO and SS/SO sentences, neither the main effect of sentence type or load nor the interaction between the two factors was significant.

Digit Recall Accuracy

Digit recall errors for only those sentences that were both plausible and accurately judged were analyzed. Each of the three sentence pairs (CS vs. CO, OS vs. SO, and SS vs. SO) were analyzed with a separate ANOVA. Figure 3 shows the mean percentages of digit recall errors for the two digit-load conditions, for each sentence-pair comparison. As was expected, there was a significant main effect of

load for all sentence pair comparisons [CS/CO, $F_1(1,35) = 29.89$, $MS_e = 4.23$, $p < .0001$; OS/SO, $F_1(1,35) = 36.67$, $MS_e = 2.24$, $p < .0001$; SS/SO, $F_1(1,35) = 25.84$, $MS_e = 2.69$, $p < .0001$]. Tukey's post hoc tests indicated that, as was expected, the participants made more errors on digit recall in the five-digit load condition than in the three-digit load condition. The main effect of sentence type and the load $\times$ sentence type interaction were not significant.

DISCUSSION

The results of these studies indicate that the AMW paradigm was sensitive to local increases in processing load. All sentence pairs, in all three conditions (no load, three digits, and five digits) showed increases in listening times

at the points in the sentence where processing demands are expected to increase due to the need to construct complex syntactic structures. The increase in processing time at the verbs in object-relativized sentences was not a simple end-of-clause effect, since listening times for these verbs were longer than those for the nouns in comparable end-of-clause positions in subject-relativized sentences. The listening time effects in SO sentences, as compared with both OS and SS sentences, were on the first verb—the verb of the embedded relative clause. This position is one of two positions at which such effects are expected; the other is on the second, main verb, as was reported by King and Just (1991). Differences in whether the effect of syntactic load is found on the embedded or the main verb of SO sentences may be due to differences in task demands. As in the CO/CS contrast, the increased listening times for the embedded verb of SO sentences was not a simple end-of-clause effect, since listening times for these verbs were longer than those for the nouns in comparable end-of-clause positions in subject-relativized (OS and SS) sentences.

The goal of this study was to investigate the effect of an extrinsic memory load on on-line syntactic processing. The failure to find a significant load $\times$ sentence type $\times$ phrase interaction indicates that on-line syntactic processing was not affected by working memory resource reductions caused by an extrinsic memory load. However, overall listening times became longer as the memory load increased, indicating that the digit load increased processing load during on-line sentence comprehension. The effects of load were the same in the simple and the complex versions of the sentences in each pair. This indicates that the digit load affected one or more operations that were common to the two sentence types in each comparison in this task. These operations range from recognizing lexical items to planning motor responses. One consideration suggests that syntactic operations shared between the sentence types are not the locus of the effect of load. The failure to find a significant load $\times$ sentence type $\times$ phrase interaction indicates that the on-line syntactic operations that *differ* between the complex and the simple sentences in these studies were not affected by these loads. These operations (processing of long-distance dependencies) are highly demanding of working memory resources, whereas the syntactic operations that are shared between the sentence types contrasted here (constructing local phrase markers) are considerably less demanding of such resources. As was noted above, the single-resource theory predicts that the effect of load will be seen to a greater extent on more demanding operations. We therefore consider it unlikely that the external load was shared with syntactic operations and conclude that this study provides evidence that on-line syntactic processing is not affected by an extrinsic memory load.

With respect to end-of-sentence measures, plausibility judgment accuracy was greater for the simpler CS sentences than for the more complex CO sentences. The failure to find effects of sentence type in the SO sentences, as compared with the OS and SS sentences, probably reflects the fact that the increased load in the SO sentences occurs

well before the sentence is over. There were no interactions between sentence complexity and load in any of the end-of-sentence measures (judgment accuracy [A' scores], judgment RT, or digit recall data) in the comparison of SO and OS or SS sentences. In the comparison of CO and CS sentences, there was a significant interaction of sentence type and load in the RT data. As Figure 2 shows, this interaction was due to longer RTs for the simpler CS sentences in the two load conditions, which is not the effect of load predicted by the single-resource theory of working memory.

In summary, the concurrent memory load in the present study increased the overall difficulty of the on-line and off-line components of the AMW task. However, syntactically complex sentences were not more affected than simple sentences by the concurrent memory load in either the on-line or the off-line data. The results of this study thus indicate that a concurrent memory load does not reduce the availability of working memory resources used for on-line syntactic processing. They thus provide evidence that the working memory system used for assigning syntactic structure is a specialized resource system. In terms of Baddeley's (1986) model of working memory, it can be considered to constitute a specialized part of the central executive. Standard working memory tasks utilize and measure other aspects of the working memory system, such as phonological storage, rehearsal, and the capacity to divide attention between tasks (Engle, Tuholski, Laughlin, & Conway, 1999). The specialized portion of the central executive for which this study provides evidence does not contribute to most of these tasks.

REFERENCES

- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.
- BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- BALOGH, J., ZURIF, E. B., PRATHER, P., SWINNEY, D., & FINKEL, L. (1998). Gap filling and end of sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain & Language*, **61**, 169-182.
- CAPLAN, D., HILDEBRANDT, N., & WATERS, G. S. (1994). Interaction of verb selectional restrictions, noun animacy and syntactic processing. *Language & Cognitive Processes*, **9**, 549-585.
- CAPLAN, D., & WATERS, G. (1999). Verbal working memory and sentence comprehension. *Behavioral & Brain Sciences*, **22**, 77-126.
- COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.
- DICK, F., BATES, E., WULFECK, B., UTMAN, J., DRONKERS, N., & GERNSBACHER, M. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review*, **108**, 759-788.
- DUNN, J. (1994). *SoundEdit* (Version 2). Micromedia Inc., Apple Inc.
- ENGLE, R. W., TUHOLSKI, S. W., LAUGHLIN, J. E., & CONWAY, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, **128**, 309-331.
- FERREIRA, F., & CLIFTON, C. (1986). The independence of syntactic processing. *Journal of Memory & Language*, **25**, 348-368.
- FERREIRA, F., HENDERSON, J., ANES, M., WEEKS, P., & MCFARLANE, D. (1996). Effects of lexical frequency and syntactic complexity in spo-

- ken language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 324-335.
- GIBSON, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, **68**, 1-76.
- JUST, M. A., & CARPENTER, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122-149.
- KING, J., & JUST, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory & Language*, **30**, 580-602.
- KUČERA, H., & FRANCIS, N. (1970). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- OSTERHOUT, L., & HOLCOMB, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory & Language*, **31**, 785-804.
- VOS, S. H., GUNTER, T. C., SCHRIEFERS, H., & FRIEDERICI, A. D. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language & Cognitive Processes*, **16**, 65-103.
- 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* (pp. 531-555). London: Erlbaum.
- WATERS, G., CAPLAN, D., & ROCHON, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology*, **12**, 1-30.

(Manuscript received September 11, 2001;
revision accepted for publication February 25, 2002.)