PROCESSING QUANTIFIER SCOPE*

Veena D. Dwivedi, Matthew Goldhawk & Patrick Wood
Brock University

1. Introduction

Recent work in language processing (Christianson, Hollingworth & Ferreira, 2001; Sanford & Sturt, 2002; Swets, Desmet, Clifton & Ferreira, 2008) suggests that interpretive processes are often incomplete, such that comprehenders do not commit to a particular meaning during a parse. Shallow interpretive processes have implications for understanding ambiguity at the syntax-semantics interface, particularly for scope ambiguous sentences, such as (1a):

(1) a. Every kid climbed a tree.
    b. The trees were in the park.
    c. The tree was in the park.

Sentences such as (1a) are ambiguous, despite the fact that they lack any syntactic or lexical ambiguity. The different meanings are the result of different logical orders in which the quantifiers are interpreted. On one interpretation, it is the case that for *every* (∀) child, a (∃) tree was climbed, which results in an inference that several trees were climbed. This reading is called the 'surface scope' reading, since the order of interpretation of the quantifiers matches the surface linear order of the quantifiers in the sentence (see 2a). On another reading, called the inverse scope reading, the interpretation is that it is the case that there is a (∃) tree, such that every (∀) kid climbed it (see 2b). The inverse scope reading results in a meaning where just one tree was climbed. The logical formulae for these interpretations are given below:

(2) a. $(\forall x)$ (x is a kid)$\rightarrow$ $(\exists y)$ (y is a tree & x climbed y)
    [read as: "For every kid x, there is a tree y, such that x climbed y"]

  b. $(\exists y)$ (y is a tree) & $(\forall x)$ (x is a kid $\rightarrow$ x climbed y)
    [read as: "There is a tree y, such that for every kid, x, x climbed y"]

In an end-of-sentence on-line grammatical acceptability task, Kurtzman & MacDonald (1993) (henceforth K&M; see also Anderson, 2004) showed that plural continuation sentences (1b), consistent with a surface scope interpretation of (1a), are preferred over singular continuations (1c), consistent with the inverse scope interpretation. From a processing perspective, the surface scope bias for (1a) is expected. That is, because that interpretation is consistent with the linear order of the sentence, as it requires the least amount of linguistic structure to

* We would like to thank Jane Dywan, Tim Murphy and Ayda Tekok-Kilic for helpful discussion, in addition to the audience at CLA. Thanks to Christine Luckhart for help in manuscript preparation. This work was partially supported by SSHRC (#410-2006-1748), as well as Brock Research Funds.

represent at the level of Logical Form (May, 1985). As such, K&M's finding is supported by the Minimal Structure Hypothesis of Dwivedi (1996) (see also Tunstall, 1998 and Anderson 2004):

(3) *Minimal Structure Hypothesis*

When constructing parse, postulate only as much structure as is required by the well-formedness rules of the grammar.

However, recent behavioural findings indicate that K&M's findings have not been fully replicated (see Tunstall, 1998; Filik, Paterson & Liversedge, 2004; Paterson, Filik & Liversedge, 2006, as well as Anderson, 2004). One potential reason why findings have been equivocal is that the above-mentioned studies examined several linguistic factors simultaneously—e.g., type of verb phrase, type of verb, type of quantifier, order of quantifiers. Moreover, Kemtes & Kemper (1999) showed that judgments for sentences like (1a) are modulated by age and verbal Working Memory (WM) span. Perhaps the lack of replication could be explained by individual differences between participant groups across studies.

In a recent Event Related Potential (ERP) study conducted by Dwivedi, Phillips, Einagel, and Baum (2010) the interpretation of sentences such as (1a) was investigated by examining neurophysiological responses to continuation sentences (1b) and (1c). All ambiguous context sentences had the same form, which was "***Every*** N^Verb^ ***a*** Direct Object" and were followed by either singular or plural continuation sentences as in (1b,c). Responses to these sentences were compared to those that occurred after Control contexts, modeled after K&M, which were "***Every*** N^Verb^***a different***/***the same*** DO." Our results were not consistent with those found by K&M; there was no neurophysiological evidence for a preference of the plural continuation. Instead, Ambiguous continuation sentences patterned together, exhibiting a late sustained negative-going ERP component 900 ms after the presentation of the Noun "tree(s)" and lasting throughout the presentation of the auxiliary Verb "was/were" (for details see Dwivedi et al., 2010). This finding was interpreted as evidence that in very early stages of linguistic analysis, the parser/brain leaves scope ambiguous sentences as "underspecified", that is, it does not commit to an interpretation of such sentences.

The lack of the replication of the K&M finding was unexpected, especially given processing principles such as the Minimal Structure Hypothesis. As a follow-up to the ERP study, Dwivedi & Goldhawk (2009a,b; henceforth D&G) conducted a self-paced reading study of scope ambiguous sentences, where the goal was simply to show an empirical basis for the plural reading (corresponding to surface scope), as a baseline experiment. In order to ensure such a finding, we pre-selected a subset of stimuli from Dwivedi et al., 2010, which also reported an off-line norming study (see Methods section below for details). Whereas that norming study reported that quantifier scope ambiguous sentences were preferentially interpreted on their surface scope reading (at about 74%), an items analysis revealed huge variability across the set of 160 stimuli tested. We thus pre-selected 24 items that were heavily biased for the plural (surface scope) interpretation and then tested how these stimuli would be perceived in an on-line experiment. Furthermore, we modified the control conditions; these now used referential determiners such as “that” and “those”, in order ensure other scope ambiguous interpretations would not ensue (see D&G, and Dwivedi et al., 2010 for discussion). The control context sentences were of the following form:

(4) a. Every kid climbed those trees.
    b. Every kid climbed that tree.

In the sentences above, (4a) ensures that the plural continuation (1b) should follow unambiguously, whereas (4b) is control context for the singular continuation (1c). The conditions are summarized in Table 1 below:

Table 1: Stimuli from Dwivedi & Goldhawk (2009a,b)

Context

		Ambiguous	**Control**
Number	**Plural continuation**	Every kid climbed a tree. The trees were in the park.	Every kid climbed those trees. The trees were in the park.
	Singular continuation	Every kid climbed a tree. The tree was in the park.	Every kid climbed that tree. The tree was in the park.

Thus, the within-subjects study was defined by two independent variables: type of Context (Ambiguous (A) or Control (C)) and type of Continuation sentence (Plural (P) or Singular (S)). We also investigated the role of WM in processing such ambiguous sentences. We predicted to find a bias consistent with off-line judgments, such that the singular continuation sentence would be dispreferred following the ambiguous context sentence. As a result, RTs for Ambiguous Singular (AS) continuation sentences should be reliably longer (as a result of the ensuing revision) than those for Control Singular (CS), where no revision would be necessary. Furthermore, RTs for Ambiguous Plural (AP) should not differ significantly from their control (CP). In addition, due to findings in Dwivedi et al., 2010, we predicted that effects would be found late in the sentence, after the verb. Regarding span effects, either it could be the case that only the high WM group (n=40) would show effects of the bias, presumably because they have the capacity to handle ambiguity, whereas the low WM group would be insensitive to ambiguity (see Fiebach et al., 2002) due to diminished capacity. Another possibility would be that both groups would be sensitive to the ambiguity, and if so, then individuals with Low WM span would have a harder time revising a dispreferred interpretation, so that increased RTs would be observed for this group and not the High WM group. In any case, we predicted an interaction between the linguistic factors and WM span.

We ran 80 participants and did not find any reliable difference in RTs for continuation sentences, in any region. Instead, what we found was an effect at the first sentence, where the final word took longer to read in Control contexts vs. Ambiguous contexts. There was a strong trend indicating that this effect was modulated by the high WM group. We interpreted these results in terms of findings by Traxler et al., (1998) and Swets et al., (2008) who found that participants spent more time reading disambiguated sentences vs. ambiguous sentences. They argued that ambiguous sentences show an RT advantage, since readers do not spend the effort to interpret them. As such, they spend less time on ambiguous sentences, due to shallow processing vs. unambiguous sentences, which are fully interpreted. D&G hypothesized that their results could be explained in a similar manner, since scope ambiguous sentences were read faster than control sentences. Because these were left as unresolved, both plural andr singular continuation sentences would be equally coherent, resulting in no difference in RTs. As such, our self-paced RT study replicated the findings of the ERP study, where no bias for plural sentences, consistent with the surface scope interpretation, was observed.

Given the evidence that scope ambiguous structures are processed in a shallow manner, the present experiment sought to modulate the depth of processing by including questions regarding scope interpretation after critical trials. That is, in D&G, superficial content questions followed only a subset of filler trials in order to ensure that participants were paying attention. In the present study, all filler trials will be followed by superficial content questions, and, importantly, all critical trials will be followed by questions regarding scope interpretation. Given that participants are now given the goal of resolving the ambiguity, the predictions outlined in D&G should hold, where AS will be the dispreferred continuation, resulting in increased RTs. Furthermore, linguistic factors of interest should interact with WM span. Finally, if shallow processing is the result of reduced cognitive capacity (Hannon & Daneman, 2001), we predict that the Low WM group will be less accurate in responding to questions that are embedded in ambiguous vs. control contexts, where the High span group should not exhibit response accuracy differences.

2.0 Methods

2.1 Participants

Forty-eight right-handed native speakers of English (39 female, mean age 20.8 years, range 18 to 30 years) were recruited at Brock University and were either paid $10 each to participate in the experiment or were given partial course credit (if applicable).

2.2 Materials

Twenty-four experimental stimuli were prepared such that each consisted of 2 sentences: a Context sentence and a Continuation sentence. The Context sentence always began with “Every NP” as a subject, and the direct object was either an NP preceded by an existential quantifier (“a”) for Ambiguous contexts, or a demonstrative determiner (“that/those”) for Control contexts. The use of referential determiners (Kaplan, 1978) would ensure that no scope ambiguity could occur with Control context sentences. Continuation sentences began with a singular or plural subject noun phrase and auxiliary verb (“The tree(s) was/were"; "The melon(s) was/were"), followed by either a prepositional phrase ("in the park") or conjoined adjectives ("soft and juicy"). Please see Table 1 above.

The 24 target sentences were combined with 64 stimuli from an unrelated experiment, and 101 fillers, for a total of 189 items. The target sentences were divided into four lists, ensuring that all factors were counterbalanced in Latin square format. Whereas in D&G, only a subset of filler trials were followed by superficial content questions, in the present experiment, all filler trials were followed by such questions. In addition, in contrast to D&G, all target sentences were now followed by questions; these were not superficial as they directly queried the interpretation (see Anderson, 2004).

(5) Every kid climbed a tree.
The trees were in the park.

How many trees were climbed?
1) SEVERAL 2) ONE

All questions were forced choice, with two buttons (labeled as "1" and "2") designated for answer selection. Participants pressed the button that corresponded to the answer on the screen. Answers were counterbalanced such that an equal number of correct answers were displayed on the right and left side of the screen.

2.2.1 Stimuli

The target sentences used were exactly the same as in D&G. The 24 ambiguous context sentences were selected from a previous off-line study reported in Dwivedi et al. (2010), where two semi-randomized lists were created and 32 subjects (none of whom participated in the present experiment) read ambiguous context sentences as above, and were asked to circle their preference (see Fig. 1). In this off-line task, discourses were presented in a booklet in a pseudo-random order, with the constraint that no more than two of the same type of trial succeeded one another. In each list, 80 ambiguous context sentences were presented, as well as 80 unambiguous ones (40 Control Singular and 40 Control Plural, as in Table 1 above). Note that plural and singular continuation choices were counterbalanced to appear either on the top or bottom position. In addition, 80 fillers were used from an unrelated experiment. Results were consistent with those of K&M, such that the plural interpretations were preferred for Ambiguous contexts such as *Every kid climbed a tree* 74% of the time. In D&G, an items analysis was conducted. Results indicated that not all items were biased in the same way, such that plural judgments ranged from 20-100%. The 24 items used in the present study (as in D&G) were 93-100% plurally biased, i.e., heavily biased for surface scope interpretation (11 items judged as plural at 100%, 9 at 94%, and 4 at 93%).

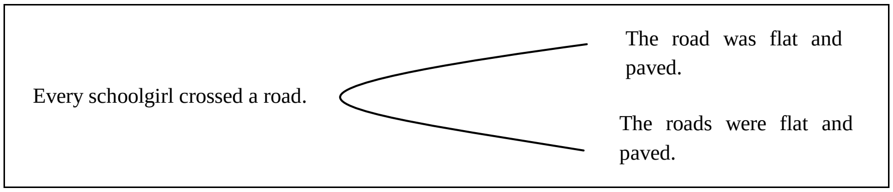

Fig. 1: An example of an ambiguous pre-test item in Dwivedi et al., (2010).

2.3 Procedure

A verbal Working Memory task derived from Daneman & Carpenter (1980) was administered to participants prior to the computer-based task (Siegel & Ryan, 1989). For reasons of space, please see Dwivedi & Goldhawk 2009b for details of the self-paced reading procedure, as well as the scoring of the verbal WM task. In the current study, the mean span was 7.25 (*SD = 1.94*; range = 3 to 12. Participants scoring higher than the mean (8 or higher) were categorized as "High WM", while participants scoring 7 or lower were placed in the "Low WM" category.

3. Results

Outlier RT data were filtered by establishing upper and lower boundary values; such that any data point (within subject, condition, and word position) exceeding 2 standard deviations in either direction was attenuated to the nearest ceiling value. This affected less than 2% of the data.

All statistical analyses reported below concern reading times recorded per word at the Context sentence (S1) and Continuation sentence (S2).

Separate repeated measure mixed ANOVAs were conducted for S1 and S2. The factors included were: WM (2 levels: High vs. Low), Context (2 levels: Ambiguous vs. Control), Number (2 levels: Plural vs. Singular), and Word Position (number of levels determined by region, reported below). Our hypotheses regarding differences between groups at regions defined below were confirmed; as a result, analyses reported below will be separated by group.[1]

The RT analyses reported below used PASW[2] v18 statistical software and employed the Greenhouse-Geisser (1959) non-sphericity correction for effects with more than one degree of freedom in the numerator. Following convention, unadjusted degrees of freedom are reported, along with the Greenhouse-Geisser epsilon value (ε) and adjusted *p*-value. Mean square error values reported are

[1]For S1, a mixed ANOVA performed over the whole sentence revealed a reliable interaction between both Context x Working Memory (F_1 (1, 46) = 4.6, *MSE* = 22 230; *p* = 0.04; F_2 (1, 46) = 9.7, *MSE* = 13 416; *p* = 0.003) and Number x Working Memory (F_1 (1, 46) = 4.7, *MSE* = 10 141; *p* = 0.03; F_2 (1, 46) = 2.3, *MSE* = 22 370; *p* = 0.14). For S2, in the post-disambiguation region (V1-V2-V3), a reliable interaction was revealed between Number and Working Memory (F_1 (1, 46) = 7.6, *MSE* = 10 243; *p* = 0.008; F_2 (1, 46) = 5.5, *MSE* = 13 321; *p* = 0.02).

[2] Formerly known as SPSS.

those corresponding to the Greenhouse-Geisser correction. All significant main effects are reported first, followed by the highest order interaction effects involving Context and/or Number.

3.1 Reading Times

3.1.1 Results for Context Sentence (S1)

Analyses at S1 were conducted over the whole sentence. Figs. 2 and 3 reveal stark differences in the pattern of reading times as exhibited by High vs. Low WM groups at the Context sentence. Namely, the High WM group exhibited a relatively flat reading rate for Ambiguous context sentences. This was very different from their behaviour reading Control contexts, where RTs increased at the end of sentence. Visual data patterns as observed for the High group were confirmed in a within-subjects 2 x 2 x 5 ANOVA involving Context (Ambiguous vs. Control), Number (Plural vs. Singular) and Word Position (5 levels: Quantifier, Subject Noun, Verb, Article, Direct Object) based on raw RT. Results revealed a strong main effect of Context (F_1 (1, 23) = 6.9, *MSE* = 33 364; p = 0.015; F_2 (1, 23) = 18.4, *MSE* = 12 703, p = 0.000), as well as a Context x Word Position interaction (F_1 (4, 92) = 4.1, *MSE* = 34 414; p = 0.03, ε = .50; F_2 (4, 92) = 6.8, *MSE* = 19 507, p = 0.003, ε = .50). Thus, Control context sentences were read more slowly by the High WM group vs. Ambiguous contexts (435 ms vs. 391 ms). Simple effects analyses revealed that this difference was strongest at the final word position, where the Control condition was read at 604 ms. in comparison to the Ambiguous condition at 469 ms. Thus, it seems that the High WM group paid more attention to the Control contexts, which used referential determiners without any previous context. Furthermore, this group showed no difficulty or complexity effects while reading scope ambiguous context sentences.

As Fig. 3 shows, the Low WM group treated Ambiguous and Control context sentences similarly, where both conditions showed increased RTs at the end of the sentence. No significant differences were revealed (all *F*s <1). Thus, for this group Control conditions were read at 400 ms in comparison to the Ambiguous conditions (403ms).

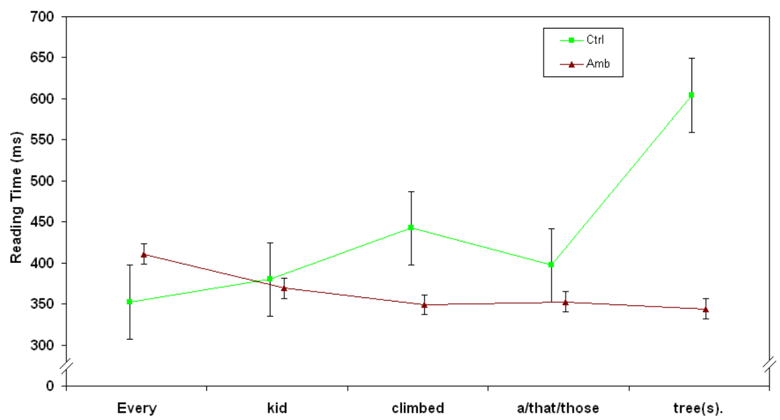

Fig. 2 Reading time in milliseconds at S1 for High WM group (*n*=24). Points represent the mean RTs per word; vertical lines depict standard error of the means.

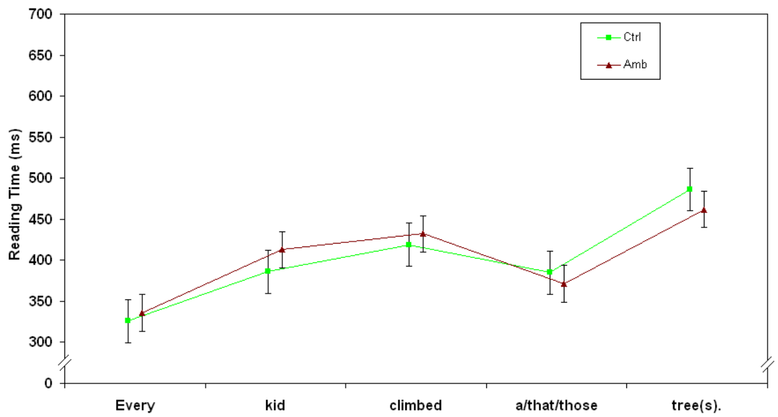

Fig. 3 Reading time in milliseconds at S1 for Low WM group (*n*=24). Points represent the mean RTs per word; vertical lines depict standard error of the means.

3.1.2 Results for Continuation Sentence (S2)

As in D&G, no reliable effects (all *Fs* <2) were revealed in the Subject-Verb region (eg, *The tree(s) was/were*…). As such, below we report results for the post-disambiguation region, after the Verb, V1-V2-V3[3] (eg., *in the park*). Examining Figs. 4 and 5, it is apparent that at the continuation sentence, now it is the Low WM group that is differentiating between conditions, in contrast to the High WM group, which does not.

That is, Fig 4 reveals that at S2, the High WM group did not differentiate between conditions. This is made clear in Fig. 6, where the mean RT for the post-disambiguation region is reported by condition. Paired-samples t-tests comparing reaction times revealed no significant difference for AS (M= 424ms, SD=91) and its control CS (M= 437ms, SD=107); t(23)= -0.54, p=.60, nor for AP (M= 448ms, SD=115) and its control CP (M= 422ms, SD=103); t(23)= 1.15, p=.26.

In contrast, for the Low WM group, Fig. 7 shows a different pattern. Paired-samples t-tests comparing reaction times do reveal a significant difference for AS (M= 470ms, SD=133) and its control CS (M= 420ms, SD=75); t(23)= 2.41, p=.025, whereas no significant difference was shown for AP (M= 409ms, SD=80) and its control CP (M= 397ms, SD=94); t(23)= .79, p=.44.

[3] Note that V3 always corresponded to the final word of the sentence.

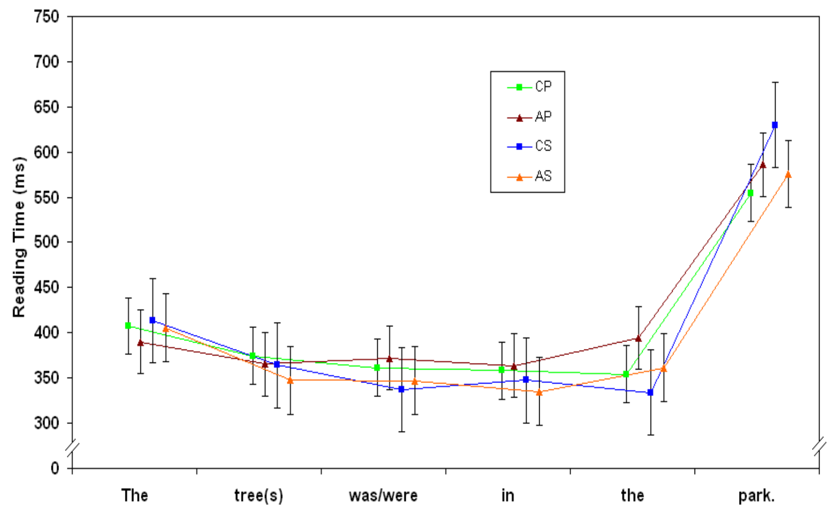

Fig. 4 Reading time in milliseconds at S2 for High WM group (*n*=24). Points represent the mean RTs per word; vertical lines depict standard error of the means.

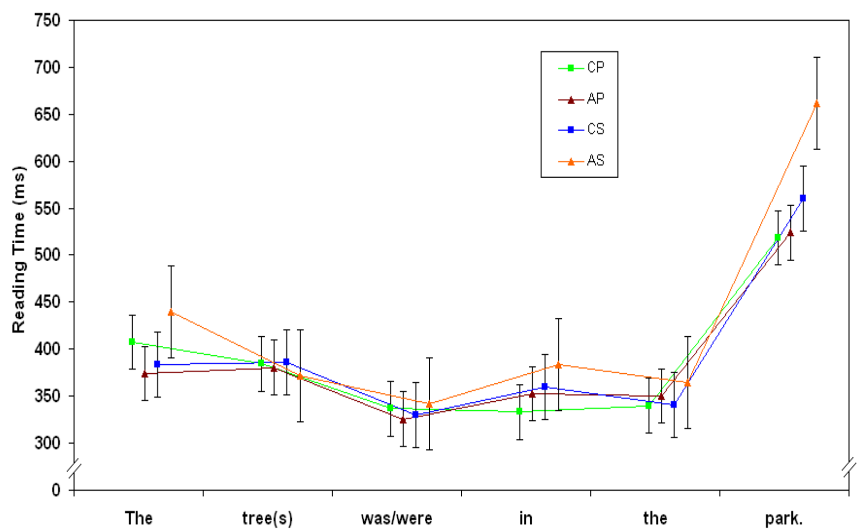

Fig. 5 Reading time in milliseconds at S2 for Low WM group (*n*=24). Points represent the mean RTs per word; vertical lines depict standard error of the means.

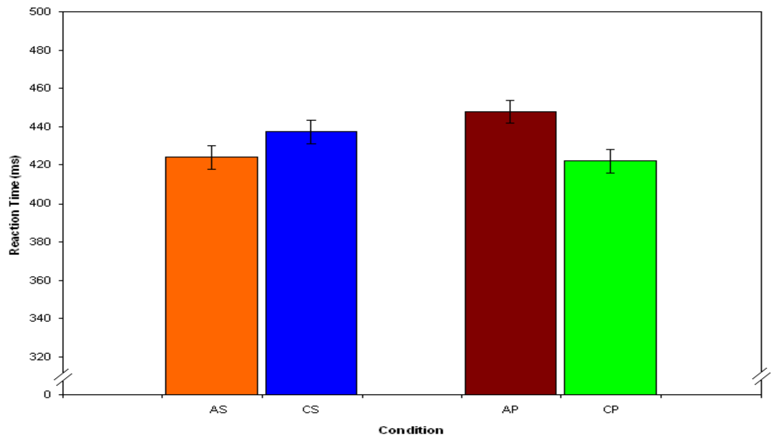

Fig. 6 Mean reading times in milliseconds at post-disambiguation region ("in-the-park") for High WM group (*n*=24)

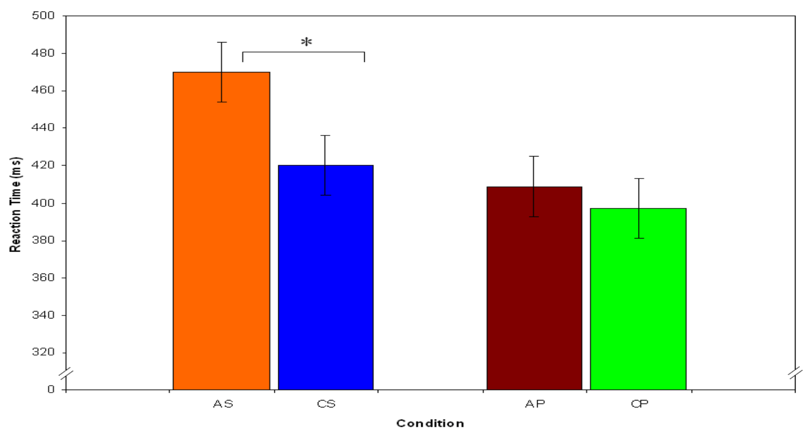

Fig. 7 Mean reading times in milliseconds at post-disambiguation region ("in-the-park") for Low WM group (*n*=24), **p*<.05

3.1.2 Question Answer Results

Our *a priori* hypothesis was that the Low WM group would show increased difficulty in responding to questions after ambiguous contexts, due to their diminished verbal WM capacity, as compared to the High WM group. Indeed, as Fig. 8 shows, the Low WM group showed difficulty in responding to questions embedded in Ambiguous vs. Control contexts. Paired-samples t-tests comparing mean question response accuracy revealed a significant difference for AS (M= .45, SD=.28) and its control CS (M= .71, SD=.25); t(23)= -4.42, p<.001, as well as for AP (M= .78, SD=.18) and its control CP (M= .94, SD=.10); t(23)= -3.94, p=.001. In other words, not only did the Low WM group have difficulty in responding to questions regarding the dispreferred

interpretation, they also had difficulty regarding queries for the preferred (plural) interpretation. In other words, the Low WM group exhibited difficulty in responding to questions after Ambiguous vs. Control contexts, as expected.
Results for the High WM group were unexpected, however. This group, like the Low WM group, also performed *at chance* when queried about the dispreferred interpretation (see Fig. 9). That is, paired-samples t-tests comparing mean question response accuracy revealed a significant difference for AS (M= .49, SD=.28) and its control CS (M= .72, SD=.27); t(23)= -4.07, p<.001. In contrast, no significant difference was found for accuracy rates regarding the preferred interpretation, as paired-samples t-tests showed (AP (M= .88, SD=.16) vs. CP (M= .94, SD=.11); t(23)= -1.5, p=.14). Thus, whereas the High WM group showed no RT evidence while reading dispreferred continuation sentences (AS condition), we see evidence of the bias in accuracy rates when directly asked about their interpretation.

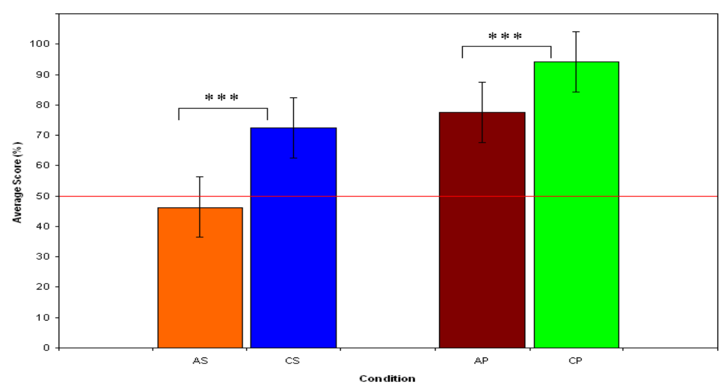

Fig. 8 Mean question response accuracy (+/- S.E) for Low WM group (*n*=24), *** *p*<.001

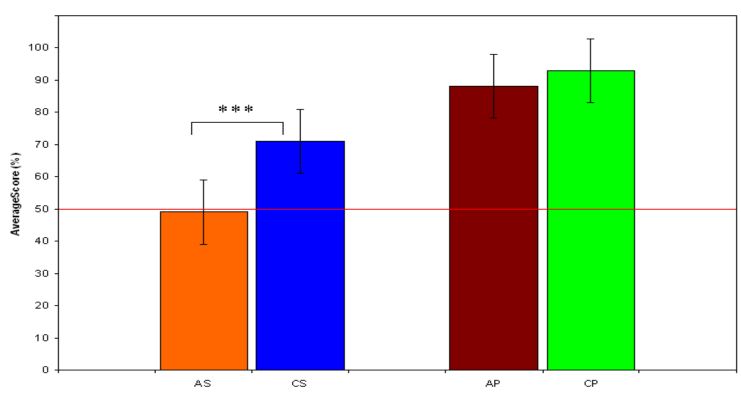

Fig. 9 Mean question response accuracy (+/- S.E) for High WM group (*n*=24), ****p*<.001

4. Discussion

The present study sought to examine the effects of task demands on the interpretation of scope ambiguous sentences, such as *Every kid climbed a tree*, across individuals that differed in terms of verbal WM span. Previously, in D&G, we examined reaction times to continuation sentences which followed scope ambiguous sentences, where the latter sentences were heavily biased towards the surface scope interpretation, consistent with the plural continuation. However, we found no RT differences between plural and singular continuation sentences following Ambiguous contexts. The only significant result of the study was that Ambiguous context sentences were read more quickly than their corresponding Controls. We interpreted these results in terms of recent findings regarding underspecification in language processing (Christianson et al., 2001; Sanford & Sturt, 2002; Swets et al., 2008). That is, our findings replicated those of Dwivedi et al., (2010) where we showed that there was no preference for the plural vs. the singular continuation sentences. Instead, the parser/brain leaves scope ambiguous sentences as underspecified constructions and simply waits for incoming material to commit to an interpretation. Nevertheless, D&G's results showing no RT differences for stimuli which are known *a priori* to be biased for a particular meaning was unexpected. In order to modulate the depth of processing, in the present experiment, we now asked participants, *How many trees were climbed?* immediately after they had read the singular or plural continuation sentence. Thus, the prediction was that participants would now fully interpret scope ambiguous context sentences, attenuating the difference found in the earlier study. Since the interpretation of scope ambiguous sentences would be complete, we also predicted to find evidence of the bias in the continuation sentence, such that AS continuations would take longer to read than CS. Furthermore, if we conceive of underspecification as a shallow processing strategy used to conserve cognitive resources, then we would expect to see the aforementioned effects in the Low WM group, since this group has diminished cognitive resources. Finally, regarding question response accuracy, we predicted that accuracy rates would be lower for the Low WM group in the Ambiguous vs Control conditions, due to the extra complexity induced by ambiguity.

Our results may be summarized as follows: First, at S1, the Context sentence, we did find an RT attenuation between Ambiguous and Control contexts, where this occurred only for the Low WM group. Second, at S2, the Continuation sentence, we did find evidence that the singular continuation was dispreferred—RTs were longer for condition AS vs. CS[4]; whereas no significant difference was found between AP and CP. Again, this difference was only found in the Low WM group. Finally, our prediction regarding increased difficulty for the Low WM group regarding question-response accuracy after ambiguous contexts was confirmed. This group performed below chance when queried about the dispreferred interpretation (AS). While accuracy rates were higher for AP (78%), this still differed reliably from its Control (94%), confirming an

[4] We note here that the lower than expected accuracy rate for CS (71%) is orthogonal to the hypotheses under investigation and leave this finding for future research in the status of singular vs. plural demonstratives (cf. King, 2001)

overall difficulty with questions that were embedded in Ambiguous vs. Control contexts for this group. We discuss the significance of these findings below.

First, we did find a difference regarding sensitivity to ambiguity between groups. That is, the Low WM group took more time to process Ambiguous context sentences, and showed evidence of the bias for the plural continuation in RTs to dispreferred continuation sentences. The High WM group did not.

As noted above, the inclusion of questions in this experiment did not attenuate the difference between Ambiguous and Control sentences for the High WM group. If anything, the reading rate for Ambiguous contexts looked flat, whereas RTs increased at the post-verbal region for the Control contexts. This lack of attention or underspecification, is also apparent at S2. Reaction times for either plural or singular interpretation did not differ from their controls. In other words, there is no empirical evidence that the High WM group is in fact doing the work of interpreting the scope ambiguous context sentences, even in an experiment where the goal of interpretation is made explicit, and stimuli are heavily biased. As such, we see no evidence of the bias in RTs at the Continuation sentence. [5] In contrast, the Low WM group did show empirical effects of the pragmatic bias for the plural interpretation at continuation sentences.

What are we to conclude from these individual differences in reading times for contexts exhibiting scope ambiguity? The key result lies in the question-response accuracy rates for both High and Low WM groups. Interestingly, here the groups did not differ; both performed at below chance when queried about the dispreferred condition, AS. Thus, although the High WM group did not show any effects of the scope (pragmatic) bias in RTs, this was evident when asked about their interpretation. These findings are reminiscent of results found in a recent self-paced reading study (Christianson et al., 2001). There, sentences such as *While Anne dressed the baby played in the crib* were examined. Participants were asked two kinds of questions. The first kind consisted of superficial content questions, such as whether it was a baby that played in the crib. This sort of question was answered with a high level of accuracy. However, when the second kind of question was asked, which was based on the misanalysis (that Anne dressed the baby), participants' accuracy rates were very low. This was the case despite the fact that, according to those authors, *dress* is a reflexive transitive verb, such that Anna dressing the baby is grammatically disallowed. Christianson et al. (2001) explain that the ungrammatical but plausible reading "lingers" in memory precisely because readers perform shallow processing in their revision. This reasoning lends itself well to the present results. That is, the High WM group does not exhibit any cost

[5] The fact that at S1, the High WM group still showed an advantage of Ambiguous context sentences can have two possible interpretations: either it is the case that interpreting ambiguous sentences is not a costly endeavour for this group; or it is the case this group was paying attention to the use of the referential determiners *that/those* in the Control sentences which did not refer to anything prior in the discourse. In other words, maybe this group does spend more time at scope ambiguous sentences but this effect is washed out by the larger amount of time spent at Control sentences. Future experiment investigating the role of previous context are planned to adjudicate between these possible explanations.

in processing scope ambiguity on-line—they have the resources to handle ambiguity, and to handle the revision required at S2 for the dispreferred AS condition. However, this shallow revision allows for the other (preferred and coherent) interpretation to linger, which explains their extremely poor performance when asked about their interpretation of the sentences.

The Low WM group, on the other hand, do not have the same level of verbal cognitive resources available to them. As a result, when they process ambiguity, there is a cost (explaining the attenuation effects apparent at S1), and when they have to revise an interpretation, again, there is a cost (explaining the longer RTs for AS). However, this revision is still shallow, and so when asked about their interpretation of the dispreferred continuation, they too perform at below chance, just like the High WM group.

In sum, we have argued that scope ambiguous sentences are constructions that the parser/brain is content to leave as only partially interpreted. That is, even when cognitive resources are available, the parser/brain does not do the work of interpreting these sentences. Between-group differences were evident in terms of on-line RTs to sentences, indicating that when task demands require the interpretation of scope ambiguous sentences, we see that individuals with Low WM capacity exhibit difficulty with these sentences, unlike the High WM group. However, queries regarding final interpretation were comparable between groups. Both Low and High WM groups showed evidence of the pragmatic bias for the plural interpretation by performing at below chance when asked about the dispreferred singular continuation. Future studies will investigate interpretation and processing of sentences that are equi-biased with respect to scope interpretation, in order to further tease apart the role that grammar and pragmatics play in scope interpretation, as well as the role of individual differences.

References

Anderson, Catherine. 2004. The structure and real-time comprehension of quantifier scope ambiguity. Unpublished doctoral dissertation, Northwestern University.

Christianson, Keil, Andrew Hollingworth, and Fernanda Ferreira. 2001. Thematic roles assigned along garden path linger. *Cognitive Psychology* 42: 368-407.

Daneman, Meredyth, and Patricia A. Carpenter. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour* 19: 450-466.

Dwivedi, Veena. 1996. Modality and discourse processing. In: Montrul, S., Kessler Robb, M. (Eds.), McGill Working Papers in Linguistics 12, Department of Linguistics, McGill University. Montreal, 17-52.

Dwivedi, Veena, and Matthew Goldhawk. 2009a. Underspecification of scope ambiguity and anaphora: Evidence from self-paced reading. Poster presentation for the 22nd annual meeting of the CUNY Sentence Processing Conference, University of California, Davis, USA.

Dwivedi, Veena, and Matthew Goldhawk. 2009b. Semantic underspecification and anaphora. In *Proceedings of the 2009 Canadian Linguistics Association Annual Conference*, ed. F. Mailhot, 15pp.

Dwivedi, Veena, Natalie Phillips, Stephanie Einagel, and Shari Baum. 2010. The neural underpinnings of linguistic ambiguity. *Brain Research* 1311: 93-109.

Fiebach, Christian J., Matthias Schlesewsky, and Angela D. Friederici. 2002. Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions*. Journal of Memory and Language* 47: 250-272.

Filik, Ruth, Kevin B. Paterson, and Simon P. Liversedge. 2004. Processing doubly quantified sentences: Evidence from eye movements. *Psychonomic Bulletin and Review* 11: 953-959.

Filik, Ruth, Anthony J. Sanford, and Hartmut Leuthold. 2008. Processing pronouns without antecedents: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 20: 1315-1326.

Hannon, Brenda, and Meredyth Daneman. 2001. Susceptibility to semantic illusions: An individual-differences perspective. *Memory & Cognition* 29: 449-461.

Kaplan, David. 1978. Dthat. In *Pragmatics: Syntax and semantics*, ed. Peter Cole. New York: Academic Press.

King, Jeffrey C. 2001. *Complex demonstratives: a quantificational account*. MIT Press.

Kemtes, Karen A., and Susan Kemper. 1999. Aging and resolution of quantifier scope effects. *The Journals of Geronotlogy* 54B : 350-360.

Kurtzman Howard S., and Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48: 243-279.

May, Robert. 1985. *Logical form: Its structure and derivation*. MIT Press.

Paterson, Kevin B., Filik, Ruth, and Liversedge, Simon P. 2006. Competition during the processing of quantifier scope ambiguities: Evidence from eye movements during reading. *Quarterly Journal of Experimental Psychology* 61: 459-473.

Sanford, Anthony J., and Patrick Sturt. 2002. Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences* 6: 382-386.

Siegel, Linda S., and Ellen B. Ryan. 1989. The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development* 60: 973-980.

Swets, Benjamin, Timothy Desmet, Charles Clifton, and Fernanda Ferreira. 2008. Underspecification of syntactic ambiguities: evidence from self-paced reading. *Memory & Cognition* 36: 201-216.

Traxler, Matthew J., Martin J. Pickering, and Charles Clifton. 1998. Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language* 39: 558-592.

Tunstall, Susanne. 1998. The interpretation of quantifiers: Semantics and Processing. Unpublished doctoral dissertation, University of Massachusetts Amherst.