Title
Noun phrase representational complexity reduces maintenance cost in working memory by increasing distinctiveness between referents

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Lam, Chi Dat (Daniel)
Xiang, Ming

Publication Date
2022

Peer reviewed

Noun phrase representational complexity reduces maintenance cost in working memory by increasing distinctiveness between referents

Chi-Dat Lam (Daniel) (dlam8@uchicago.edu)
Department of Linguistics, 1115 E. 58th Street
Chicago, IL 60637 USA

Ming Xiang (mxiang@uchicago.edu)
Department of Linguistics, 1115 E. 58th Street
Chicago, IL 60637 USA

Abstract

Previous studies have shown that representationally complex referents are encoded slower into working memory (WM) but are retrieved faster (Hofmeister, 2011; Karimi & Ferreira, 2016). However, the cost of maintaining complex representations is still not well understood. Through two self-paced reading experiments, we investigated the cost of encoding, maintaining and retrieving complex representations in WM. While we replicated the facilitatory effect during retrieval, the slowdown during encoding was not consistent across our experiments. More critically, for the first time, our experiments demonstrated that maintaining complex representations in WM is less costly than maintaining their simple counterparts. Furthermore, we found that WM maintenance cost is reduced because complex target noun phrases are more distinct from other competing referents in WM than simple ones. Overall, our results showed that the semantic elaboration of complex representations can reduce maintenance cost and provided new perspectives into this understudied WM process.

Keywords: representational complexity; working memory; maintenance; encoding; retrieval

Background

Working memory processes - encoding, maintenance and retrieval - are essential for sentence comprehension, especially for understanding long-distance dependencies, such as those present in relative clauses, pronoun resolution and other complex constructions. For example, consider the following sentence which contains an object-extracted relative clause: ***The students*** *who Priyanka wholeheartedly* ***praised*** *submitted the report*. There are multiple syntactic dependencies in this sentence, one of which is the long-distance dependency between the noun phrase (NP) *the student* in the matrix clause and the embedded verb *praised*. In order to understand (1), comprehenders have to encode the matrix NP *the students* into WM, maintain it for a period of time until the other end of the dependency, i.e. the verb *praised*, appears. At the verb *praised*, which is looking for an object, the matrix NP *the students* is retrieved and integrated into the object position.

There has been a large body of previous work investigating the memory retrieval mechanism during online sentence processing (Dillon, Mishler, Sloggett, & Phillips, 2013; Jäger, Engelmann, & Vasishth, 2017; Lewis & Vasishth, 2005; Van Dyke & McElree, 2006; Wagers, Lau, & Phillips, 2009; Xiang, Dillon, & Phillips, 2009). Comparatively speaking, the mechanisms of WM encoding and maintenance in sentence processing are less explored. There is some evidence that memory encoding contributes to retrieval cost due to interference (Barker, Nicol, & Garrett, 2001; Gordon, Hendrick, & Johnson, 2001; Hofmeister & Vasishth, 2014; Kush, Johns, & Van Dyke, 2015; Villata, Tabor, & Franck, 2018), but few studies have addressed whether there is cost associated with encoding itself (Hofmeister, 2011; Hofmeister & Vasishth, 2014). Maintenance cost is mainly discussed in the EEG literature, especially with respect to a component called the sustained anterior negativity (Fiebach, Schlesewsky, & Friederici, 2002; J. W. King & Kutas, 1995; Phillips, Kazanina, & Abada, 2005). A recent study by Ristic, Mancini, Molinaro, and Staub (2021) also reported maintenance cost using eye tracking. These studies showed that maintaining a filler in a filler-gap dependency in WM results in an increase in processing cost, typified by either a sustained ERP signal or increased go-past reading times. It is worth noting however that the status of sustained anterior negativity is under debate (Lau, 2018).

In the current work, we seek to provide novel evidence bearing upon the memory maintenance mechanism during sentence processing. Our findings also have implications for memory encoding and retrieval. The particular empirical ground we investigate is how the representational complexity can facilitate or hinder different working memory processes. A linguistic representation is more complex than another if it has more syntactic structures and/or semantic features. A number of previous studies have shown that the representational complexity of an NP referent can modulate processing cost (Hofmeister, 2011; Hofmeister & Vasishth, 2014; Karimi, Diaz, & Ferreira, 2019; Karimi, Diaz, & Wittenberg, 2020; Karimi & Ferreira, 2016; Karimi, Swaab, & Ferreira, 2018; Troyer, Hofmeister, & Kutas, 2016). The following pair of examples were tested in Hofmeister (2011). The NP *an alleged Venezuelan communist* is more representationally complex than the NP *a communist*.

(1) It was **a communist** who the members of the club <u>banned</u> from ever entering the premises.

(2) It was **an alleged Venezuelan communist** who the members of the club <u>banned</u> from ever entering the premises.

For sentences like (1) and (2) with a simple and complex matrix NP, respectively, Hofmeister (2011) showed that reading time (RT) on *communist* is slower in (2) than in (1), but

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

on the word right after the verb *banned*, the RTs are faster in (2) than (1). This was taken as evidence that encoding a complex NP is more costly than encoding a simple one, but a more effortful encoding can facilitate memory retrieval later at the verb when the NP referent needs to be retrieved and integrated (Hofmeister, 2011; Hofmeister & Vasishth, 2014). There are a number of possible reasons to explain this effect. While encoding a referent with more features demands more resources and therefore higher cost, the resulting representation could be made more salient due to its richness in features, easing retrieval effort. The featural richness of the more representationally complex referent might also allow it to be more distinct from other competing referents, thus preventing retrieval interference. Another possibility, discussed in Karimi et al. (2020), is that because more time could be spent on encoding complex NPs, this could entail more attentional resources dedicated to complex NP referents, leading to faster retrieval.

In two experiments, we sought to 1) conceptually replicate the reported effects of target NP complexity during WM encoding and retrieval and 2) explore the cost of maintaining representationally complex NPs. We hypothesized that similar to retrieving complex representations, maintaining complex representations is also less costly. We also investigated what accounted for the difference between the cost of maintaing complex and simple target NPs. To observe WM maintenance of target NPs, Experiments 1 and 2 employed a maintenance window between the encoding and retrieval sites. By looking at RT differences in the maintenance window, we would be able to study the cost of maintaining complex representations and understand which mechanisms account for differences in processing costs during WM maintenance.

Experiment 1

Participants

101 participants, recruited through the data collection website Prolific, participated in this experiment for payment. Sample size was determined through a power analysis of the data from a smaller pilot of 20 participants. All participants self-identified as native, monolingual English speakers who were raised in monolingual households. In addition, participants had no language related disorders or literacy difficulties. Data from 17 out of 101 participants was removed because they scored below the threshold of 75% for the comprehension questions (1 standard deviation below the mean accuracy). The experiment took about 20 minutes and participants were compensated $3.50.

Methods and Materials

The self-paced reading experiment had a 2 x 2 design. It consisted of 32 4-condition items which were manipulated in terms of the complexity of the matrix NP (Simple vs. Complex) and the type of relative clause (RC) involved (subject-extracted (SRC) vs. object-extracted (ORC)). An example is given in (3)-(6), with the slashes indicating the self-paced reading regions. In the complex NP condition, the matrix subject NP contains two prenominal modifiers (, whereas in the simple NP condition, the matrix subject NP does not contain any prenominal modifier.

All sentences consisted of the matrix subject NP, which was followed by the RC modifying the subject NP, then an adverb of time, the matrix verb and the matrix object NP. The matrix subject NP always started with *those* and was modified by either 0 (simple) or 2 (complex) nouns or adjectives. In both types of RC manipulations, the RC verb was preceded by an adverb to allow for an extended maintenance window. Full sets of experimental stimuli for both experiments can be found at `https://osf.io/7heg5/`. Example sentences for Experiment 1 are as follows:

(3) **Complex, SRC:** Those / emotional / crash / survivors / who / dutifully / assisted / Sophia / last week / joined / the meeting.

(4) **Complex, ORC:** Those / emotional / crash / survivors / who / Sophia / dutifully / assisted / last week / joined / the meeting.

(5) **Simple, SRC:** Those / survivors / who / dutifully / assisted / Sophia / last week / joined / the meeting.

(6) **Simple, ORC:** Those / survivors / who / Sophia / dutifully / assisted / last week / joined / the meeting.

Each participant read only one condition per item, totaling 32 experimental sentences. In addition, they also read 32 filler sentences, which were sentences of various types that contained long distance dependencies. Each sentence, experimental or filler, preceded a yes-no comprehension question targeting the dependency between the RC verb *assisted* and either the matrix subject NP, e.g. *those emotional crash survivors/those survivors* or the NP introduced in the RC, e.g. *Sophia*. An example comprehension question is *Was it those survivors who were assisted by Sophia?*. The expected answer for half of the questions was *Yes* and for the other half was *No*. Participants did not receive feedback on the accuracy of their answers. In this experiment, average accuracy across all items (including fillers) was 85.9%, and the average accuracy on the experimental sentences was 84.0%.

The experiment was carried out on Ibex Farm (Drummond, 2013), where the experimental and filler items were randomized and presented to participants. Participants did two practice trials before reading the experimental and filler items. Before a trial started, a dash line appeared in the middle of the screen where the stimuli would appear. Upon pressing the space bar, the dash line disappeared and the first word appeared. Participants were instructed to press the space bar to continue reading the sentence. As the space bar was pressed, the current word(s) was replaced by the subsequent word(s).

For statistical analysis, after excluding data from participants who did not meet the comprehension accuracy threshold of 75%, raw RT beyond three standard deviations of the mean raw RT at each sentence position and condition are excluded. We rejected 1.03% of the raw RTs through this procedure. Following Hofmeister and Vasishth (2014), we did not exclude RTs of sentences whose comprehension question was answered incorrectly. This was done to ensure we did not discard instances that might have reflected failure to maintain the correct dependency in WM.

RTs in each region were log-transformed and then residualized on two predictors: the linear position of a region in a sentence and log RT of the region immediately prior to the current one. Both predictors are known to impact self-paced reading RTs for independent reasons. Residualization was done using linear mixed models through the *lmer* package (Bates, Mächler, Bolker, & Walker, 2015). RTs from fillers were included in the residualization process. Residualized log RTs served as the dependent variable for our analyses.

For data analysis, we considered three sites of interest - encoding, maintenance and retrieval sites. The **encoding site** included the matrix subject NP, with the critical word being the head noun, i.e. *survivors*. RTs at this region reflected processing effort to encode the subject NP into information to be stored in WM. The **retrieval site** included only the RC verb *assisted*, where the matrix subject NP is retrieved from WM to serve as the subject or object of the RC verb. We also analyzed the spill-over region after the RC verb, the adverb of time *last week*. The **maintenance site** included the words between the encoding and retrieval sites, the complementizer *who*, the RC subject NP *Sophia* in ORCs, and the preverbal adverb *dutifully*. During this period, the matrix subject NP is maintained in WM, awaiting retrieval.

For statistical analyses, we employed Bayesian hierarchical modeling, using the R package *brms* (Bürkner, 2017). For each self-paced reading region examined, the model used 4 chains, with 2000 samples per chain, the initial 1000 samples being warm-up samples and no thinning. This led to 4000 post-warmup samples for each parameter estimate per region. Models of regions up to (and including) the complementizer *who* comprised of the fixed effect of matrix NP complexity (sum coded, simple -0.5, complex 0.5). Models of regions after the complementizer *who* comprised of fixed effects of matrix NP complexity (sum coded, simple -0.5, complex 0.5), RC type (sum coded, SRC -0.5, ORC 0.5) and their interaction. All fixed effects were sum-coded. All models also contained by-participant and by-item random intercept adjustments and random slopes for all fixed effects analyzed in that region. We used relatively weak, uninformative priors for all parameters. For the prior for all the fixed effects, including the intercept, we used a normal distribution N(0,10) with mean 0 and standard deviation of 10[1]. The final results were reported in terms of the mean of the posterior distributions and the 95% credible intervals. We considered a predictor as reliable if the credible interval does not include 0.

Results

The mean comprehension question accuracies by condition are as follows: SRC Complex: 92.1% (SD = 2.70%); SRC Simple: 90.2% (SD = 2.98%); ORC Complex: 87.5% (SD = 3.31%); ORC Simple: 85.9% (SD = 3.49%).

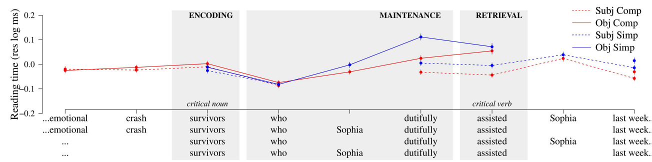

Figure 1: Residualized log reading times for Experiment 1

In the **encoding site**, as shown in Figure 2 and Table 1, there is no effect of NP complexity on the noun *survivors*, contrary to the slowdown for more complex NPs observed by Hofmeister (2011) and Hofmeister and Vasishth (2014).

In the **maintenance site**, there is no effect of NP complexity on the complementizer *who*. On the preverbal adverb *dutifully*, we observed significant effects of matrix NP complexity and RC type, as well as weak evidence for an interaction. More specifically, participants read this word faster if the matrix NP is complex. They also read faster if they were reading an SRC rather than an ORC. The weak evidence for an interaction arose from the fact that the facilitation effect due to NP complexity is more pronounced in ORCs than SRCs.

Region	Effect	Mean	Lower CrI	Upper CrI
survivors	NP	0.015	-0.021	0.048
who	NP	0.002	-0.023	0.027
dutifully	NP	-0.062	-0.087	-0.036
	RC	0.079	0.043	0.114
	NP x RC	-0.045	-0.092	0.002
assisted	NP	-0.027	-0.058	0.003
	RC	0.087	0.053	0.122
	NP x RC	0.024	-0.036	0.085
last week	NP	-0.044	-0.070	-0.020
	RC	0.030	0.005	0.057
	NP x RC	-0.002	-0.052	0.047

Table 1: Model estimates for Experiment 1 (NP: matrix NP complexity, complex vs simple; RC: RC type, ORC vs SRC)

In the **retrieval site**, on the RC verb *assisted*, only the fixed effect of RC type is significant. Similar to the maintenance region, RTs were faster in an SRC than in an ORC, which is an expected effect given previous works on the SRC advantage in English (Gibson, 1998; Gibson et al., 2000; J. King & Just, 1991; Staub, 2010; Traxler, Morris, & Seely, 2002). The RC type effect continued to be significant on the adverb *last week*. More critically, we only observed weak evidence of a speed up due to matrix NP complexity on the RC verb

[1]An example of a *brm* model:
brm(formula = residualizedlogrt ∼ npcomplexity * rctype
+ (1 + npcomplexity * rctype|participant)
+ (1 + npcomplexity * rctype|item),
data = rt_data, family = gaussian(),
prior = c(prior('normal(0,10)', class = 'Intercept'),
set_prior('normal(0,10)', class = 'sigma'),
set_prior('normal(0,10)', class = 'b'),
set_prior('normal(0,10)', class = 'sd'),
set_prior('lkj(2)', class = 'cor')),
warmup = 1000, iter = 2000, chains = 4,
control = list(adapt_delta = 0.99, max_treedepth = 12))

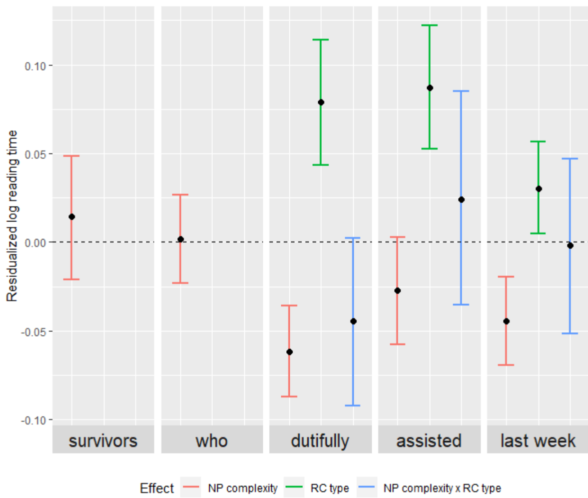

Figure 2: Credible intervals for Experiment 1

assisted. However, this speed up was reliable on the spillover region, the adverb *last week*. There was no evidence of a main effect on the matrix verb *joined* and its spillover.

Discussion

We did not replicate the slowdown in the encoding site due to matrix NP complexity, as previously observed (Hofmeister, 2011; Hofmeister & Vasishth, 2014). However, in the retrieval site, consistent with previous work (Hofmeister, 2011; Hofmeister & Vasishth, 2014; Karimi et al., 2020; Karimi & Ferreira, 2016), there was evidence in our results showing that having a complex matrix NP speeds up RT.

More importantly, a novel finding from the current work is that having a more complex matrix NP also benefited the maintenance process, as evident by the facilitation effect observed on the preverbal adverb *dutifully*. There was also some evidence that this effect was more pronounced in ORCs than in SRCs. The difference between the two clause types invited a hypothesis about the deeper mechanism based on which representational complexity can facilitate maintenance. In particular, when there are multiple competing referents that need to be maintained in WM, having a rich set of features on one referent helps keeping it apart from other referents. In SRCs, prior to the RC verb, the matrix NP *survivors* was the only available referent in WM. In ORCs, however, there were two competing referents in WM prior to the RC verb: the matrix NP *survivors* and the RC subject *Sophia*. Therefore, in ORCs, making one NP semantically rich (i.e. *the emotional crash survivors*) helps maintaining the distinctions between the two referent representations. In SRCs, since there are no competing referents, the facilitation effect thanks to the more complex NP is less pronounced. But it is important to note even though it was a smaller effect, having a complex matrix NP still resulted in faster RT in the maintenance site even for SRCs (β = -0.039, 95% CrI [-0.043, -0.008]). This implicated that matrix NP complexity also confers additional benefits during maintenance other than enhancing distinctiveness between referents in WM.

To test the hypothesis that featural richness helps maintain distinctiveness between referents, in Experiment 2, we added an additional level of embedding before the RC verb to all experimental sentences. This was done so that by the preverbal adverb *dutifully*, for both SRCs and ORCs, there are multiple competing referents that need to be maintained in WM in both SRCs and ORCs. We expect to see a larger facilitation effect for SRCs during the maintenance period due to matrix NP complexity, potentially reducing (or even eliminating) its difference from the facilitation effect observed for ORCs.

Experiment 2

Participants

100 participants, recruited through the data collection website Prolific, participated in this experiment for payment. Participant recruitment procedure was identical as Experiment 1. Data from 17 out of 100 participants was removed because they scored below the threshold of 65% (1 standard deviation below the mean accuracy).

Methods and Materials

The experiment had the same setup and materials as Experiment 1, except for the fact that an additional level of embedding *Jennifer thinks* was added after the complementizer *who* for all four conditions. Thus, Experiment 2 had the same 2 (NP complexity, sum coded, simple -0.5, complex 0.5) x 2 (RC type, sum coded, SRC -0.5, ORC 0.5) design. Example sentences for Experiment 2 are as follows:

(7) **Complex, SRC:** Those / emotional / crash / survivors / who / Jennifer / thinks / dutifully / assisted / Sophia / last week / joined / the meeting.

(8) **Complex, ORC:** Those / emotional / crash / survivors / who / Jennifer / thinks / Sophia / dutifully / assisted / last week / joined / the meeting.

(9) **Simple, SRC:** Those / survivors / who / Jennifer / thinks / dutifully / assisted / Sophia / last week / joined / the meeting.

(10) **Simple, ORC:** Those / survivors / who / Jennifer / thinks / Sophia / dutifully / assisted / last week / joined / the meeting.

The additional level of embedding was also added to some fillers and all comprehension questions. An example comprehension question is *Was it those survivors who Jennifer thinks were assisted by Sophia?*. The sentences were presented in a self-paced reading task, like in Experiment 1. In this experiment, average accuracy from all participants for the comprehension questions, including those following filler sentences, was 77.9%, and the average accuracy on the experimental sentences was 73.4%.

All statistical analyses followed the same procedure as Experiment 1. We removed 1.24% of RT data that were 3 standard deviations of the mean raw RT at each sentence position and condition. The encoding and retrieval sites were the same as Experiment 1. The maintenance site was extended to include the new level of embedding *Jennifer thinks*.

Results

The mean comprehension question accuracies by condition are as follows: SRC Complex: 82.8% (SD = 3.77%); SRC Simple: 80.0% (SD = 4.01%); ORC Complex: 73.3% (SD = 4.42%); ORC Simple: 73.2% (SD = 4.43%).

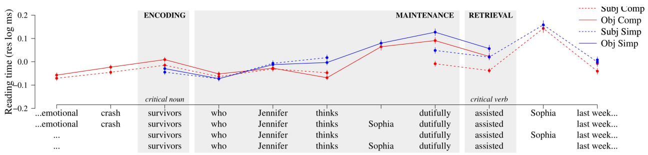

Figure 3: Residualized log reading times for Experiment 2

In the **encoding site**, as demonstrated by Figure 4 and Table 2, on the head noun *survivors*, a main effect of NP complexity was observed. More specifically, RTs were slower when the matrix NP was complex. This result was in line with the slowdown found by Hofmeister (2011) and Hofmeister and Vasishth (2014) but was different from Experiment 1.

In the **maintenance site**, there was no main effect of NP complexity on the words *who* and *Jennifer*. However, on the following verb *thinks*, there was a main effect of NP complexity, where participants read faster if the head noun was complex. It is noteworthy because this verb is the first word where there were two competing referents in WM, *survivors* and *Jennifer*. On the adverb *dutifully*, we found a main effect of NP complexity, where RTs were faster in sentences with a complex matrix NP. We also found a main effect of RC type, where sentences with an SRC were read faster in this region. Unlike Experiment 1, no evidence for an interaction between NP complexity and RC type was found. At this word, there were still two competing NP referents in SRCs (*survivors* and *Jennifer*) while there were three competing NP referents in ORCs (*survivors*, *Jennifer* and *Sophia*).

In the **retrieval site**, on the RC verb *assisted*, contrary to Experiment 1, we found a main effect of NP complexity, where RTs were faster in sentences with a complex matrix NP. This was consistent with previous results in this region shown by Hofmeister (2011) and Hofmeister and Vasishth (2014). There was weak evidence that this facilitation effect extended to the spillover region, the adverb *last week*. There was also the expected RC type effect on the RC verb *assisted*, where sentences with SRCs were read more quickly. The main effect of RC type did not persist to the adverb *last week*. There was no evidence of a main effect on the matrix verb *joined* and its spillover.

Discussion

Contrary to Experiment 1 and consistent with results from Hofmeister (2011), we found a slowdown due to matrix NP complexity in the encoding site. It is unclear why there was a divergence between the two experiments in this site since the materials up to the word *survivors* were the same between the two experiments. This difference suggests that the slowdown during encoding might be less robust and might differ

Region	Effect	Mean	Lower CrI	Upper CrI
survivors	NP	0.034	0.008	0.062
who	NP	0.015	-0.011	0.041
Jennifer	NP	-0.020	-0.045	0.006
thinks	NP	-0.065	-0.092	-0.039
dutifully	NP	-0.047	-0.076	-0.018
	RC	0.090	0.057	0.123
	NP x RC	0.020	-0.038	0.081
assisted	NP	-0.044	-0.072	-0.017
	RC	0.051	0.014	0.088
	NP x RC	0.027	-0.031	0.084
last week	NP	-0.027	-0.054	0.0003
	RC	0.021	-0.016	0.056
	NP x RC	0.025	-0.031	0.082

Table 2: Model estimates for Experiment 2

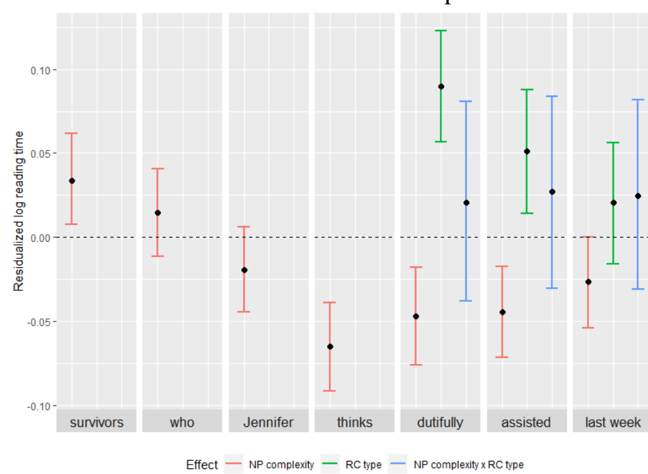

Figure 4: Credible intervals for Experiment 2

between subject pools. On the other hand, in the retrieval site, we again replicated the facilitation effect due to matrix NP complexity on the RC verb *assisted* and to some extent, the spillover region *last week*.

Critically, Experiment 2 confirmed the hypothesis that matrix NP complexity facilitates WM maintenance by allowing representations of competing referents to be distinct from each other. As noted, the verb *thinks* followed two competing referents in all conditions: *survivors* and *Jennifer*. For both SRC and ORC, the condition with a complex matrix NP was less costly on this word than the condition with a simple NP. Furthermore, the main effect of NP complexity continued on the adverb *dutifully*, and the interaction between matrix NP complexity and RC type found in Experiment 1 was absent. These results are compatible with our hypothesis: since for both SRCs and ORCs there are multiple competing referents during the maintenance period, the benefit of having distinct representations could be observed in both types of clauses.

General Discussion

In two self-paced reading experiments, we investigated the benefit and cost of encoding, maintenance and retrieval of complex linguistic representations, with a particular focus on maintenance, which has been under-explored in the sentence processing literature. For the encoding effect, we found that only one of the two experiments replicated the slowdown due to matrix NP complexity found in Hofmeister (2011), despite the two experiments having the same material during the encoding period. The inconsistency between our two experiments suggests this effect might not be robust and might be subjected to inter-participant variation. On the other hand, the previously reported facilitatory effect during retrieval due to matrix NP complexity (Hofmeister, 2011; Hofmeister & Vasishth, 2014; Karimi et al., 2020) was replicated in both experiments, on the RC verb and the spillover region.

Most importantly, our results showed, for the first time, a facilitatory effect in the maintenance site when participants read sentences with a complex matrix subject NP, suggesting a lower cost of maintaining featurally rich representations. We hypothesized that at least one possible source of this effect is due to the fact that rich semantic features on a representation helps to keep competing referent representations distinct from each other. Evidence supporting this hypothesis comes from SRCs and its comparison with ORCs. In Experiment 1, when only the ORC but not the SRC construction contained competing referents in the maintenance site, NP complexity had a smaller facilitatory effect in SRCs than ORCs; but both constructions showed similar facilitatory effects in Experiment 2 when an additional competing referent was added to the maintenance site.

We also observed that enhancing distinctness among competing referents is likely not the only reason that complex NP representations can facilitate the maintenance period. As we noted earlier, in Experiment 1, even though there was a difference between SRCs and ORCs in terms of the facilitation effect brought about by the complex NP, there was nonetheless a reliable effect in SRCs. This effect goes beyond what our proposal can account for. One possible explanation for this is the time-dependent attention account set forth by Karimi et al. (2020). In this account, there is more time devoted to encoding complex representations, making these representations more salient due to larger amount of attentional resources dedicated to them. As a result maintaining a salient complex NP could be less effortful.

In addition to the main hypothesis we proposed for the maintenance period, we also consider other possible interpretations of our results during the maintenance site. Firstly, the "maintenance" effect could have resulted from preemptive retrieval of the target NP in anticipation of the upcoming RC verb. In particular, on the adverb *dutifully*, people may have expected that a verb is coming up, and this may have triggered a retrieval of the head noun. A potential challenge for this account is that the effects observed on the adverb *dutifully* were not entirely identical as the effects observed on the actual retrieval verb *assisted*, especially in Experiment 1.

Another possible interpretation is that our "maintenance" effect might actually be the result of memory encoding. Keeping referent representations distinct from each other can be beneficial for the encoding process. In fact, when similar items are encoded in WM, encoding interference could arise (Barker et al., 2001; Gordon et al., 2001; Hofmeister & Vasishth, 2014; Kush et al., 2015; Villata et al., 2018), due to the fact that similar items compete for the shared features and result in degraded representations of one or all items, a process known as feature overwriting (Nairne, 1990; Oberauer & Kliegl, 2006). In the current case, while the rich semantic features on the matrix subject NP can give rise to higher encoding cost of this referent, subsequent encoding on later referents such as *Sophia* or *Jennifer* could be facilitated since the matrix subject NP is more distinct from them. The challenge with this account is that it should have predicted predicted the facilitation effect to arise on the proper name *Sophia* or *Jennifer*, at the moment when the encoding of the new referents took place. But in our results the facilitation effect thanks to matrix NP complexity appeared after the proper names (e.g. on *dutifully* or *thinks*). The fact that the effect appeared after the encoding of the new referents lends some support to our hypothesis that the cost of maintaining referent representations (that have already been encoded) could be reduced when competing referents have more distinct features.

Yet another different interpretation of our result is that participants might have read faster in the maintenance regions in sentences with complex matrix NP because they want to get through the maintenance region quickly to "unload" the heavier WM load. Van Dyke and McElree (2006) found that in a dual-task setup, participants tend to speed up when reading a sentence when they also have to maintain words in WM for a separate recall task. Nicenboim, Vasishth, Gattei, Sigman, and Kliegl (2015) also showed faster reading time when the WM demand increases for participants with low WM capacity. In our experiments, it is possible that participants engaged in a "good-enough" strategy during the maintenance region when faced with higher WM load of complex NPs (Ferreira & Patson, 2007). More future work is needed to futher examine this possibility, but we note that a shallower processing of the complex condition may predict worse comprehension accuracy on sentences with complex vs. simple NPs, which was not borne out in the current data.

Conclusion

In two experiments, we examined how the encoding, maintenance and retrieval of representationally complex NPs in WM differ from those of their simplex counterparts, with a special focus on the understudied maintenance period. We found that storing more complex NPs reduces maintenance cost and that one possible source of this effect can be attributed to the increased distinctiveness between the target NP and competing NP(s) in WM. Our results also replicated previous work that showed facilitatory effect of complex representation on memory retrieval. However, we only found mixed evidence that representational complexity slows down memory encoding.

Acknowledgments

This work was supported by NSF Doctoral Dissertation Research Improvement Grant, BCS-2116952. We are grateful

to our five reviewers for their helpful comments.

References

Barker, J., Nicol, J., & Garrett, M. (2001). Semantic factors in the production of number agreement. *Journal of Psycholinguistic Research*, *30*(1), 91–114.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, *80*(1), 1–28.

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103.

Drummond, A. (2013). Ibex farm. *Online server: http://spellout. net/ibexfarm*.

Ferreira, F., & Patson, N. D. (2007). The 'good enough'approach to language comprehension. *Language and linguistics compass*, *1*(1-2), 71–83.

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.

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

Gibson, E., et al. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain*, *2000*, 95–126.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of experimental psychology: learning, memory, and cognition*, *27*(6), 1411.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and cognitive processes*, *26*(3), 376–405.

Hofmeister, P., & Vasishth, S. (2014). Distinctiveness and encoding effects in online sentence comprehension. *Frontiers in psychology*, *5*, 1237.

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316–339.

Karimi, H., Diaz, M., & Ferreira, F. (2019). "a cruel king" is not the same as "a king who is cruel": Modifier position affects how words are encoded and retrieved from memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(11), 2010.

Karimi, H., Diaz, M. T., & Wittenberg, E. (2020). Sheer time spent expecting or maintaining a representation facilitates subsequent retrieval during sentence processing. In *Cogsci*.

Karimi, H., & Ferreira, F. (2016). Good-enough linguistic representations and online cognitive equilibrium in language processing. *Quarterly journal of experimental psychology*, *69*(5), 1013–1040.

Karimi, H., Swaab, T. Y., & Ferreira, F. (2018). Electrophysiological evidence for an independent effect of memory retrieval on referential processing. *Journal of Memory and Language*, *102*, 68–82.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of memory and language*, *30*(5), 580–602.

King, J. W., & Kutas, M. (1995). Who did what and 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.

Kush, D., Johns, C. L., & Van Dyke, J. A. (2015). Identifying the role of phonology in sentence-level reading. *Journal of memory and language*, *79*, 18–29.

Lau, E. (2018). Neural indices of structured sentence representation: State of the art. *Psychology of learning and motivation*, *68*, 117–142.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science*, *29*(3), 375–419.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*(3), 251–269.

Nicenboim, B., Vasishth, S., Gattei, C., Sigman, M., & Kliegl, R. (2015). Working memory differences in long-distance dependency resolution. *Frontiers in Psychology*, *6*, 312.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of memory and language*, *55*(4), 601–626.

Phillips, C., Kazanina, N., & Abada, S. H. (2005). Erp effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428.

Ristic, B., Mancini, S., Molinaro, N., & Staub, A. (2021). Maintenance cost in the processing of subject–verb dependencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*(1), 71–86.

Tanenhaus, M. K., Carlson, G. N., & Seidenberg, M. S. (1985). Do listeners compute linguistic representations. *Natural language parsing: Psychological, computational, and theoretical perspectives*, 359–408.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of memory and language*, *47*(1), 69–90.

Troyer, M., Hofmeister, P., & Kutas, M. (2016). Elaboration over a discourse facilitates retrieval in sentence processing. *Frontiers in psychology*, *7*, 374.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of memory and language*, *55*(2), 157–166.

Villata, S., Tabor, W., & Franck, J. (2018). Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in psychology*, *9*, 2.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of memory and language*, *61*(2), 206–237.

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: Erp evidence. *Brain and Language*, *108*(1), 40–55.