Studies in Second Language Acquisition (2023), **45**, 264–275
doi:10.1017/S0272263122000109

RESEARCH REPORT

Text recall and use of advance organisers in first and second language

Sara Dhaene and Evy Woumans*

Ghent University, Gent, Belgium
*Corresponding author. E-mail: evy.woumans@ugent.be

(Received 13 July 2020; Revised 14 February 2022; Accepted 17 February 2022)

Abstract

Previous research identified that studying texts in a second language (L2) as opposed to the first (L1) results in substantially weaker recall. We hypothesized that use of advance organizers (AOs) might attenuate this L2 recall cost by supporting L2 users in the construction of more solid memory representations. One hundred Dutch-English bilinguals studied two texts in either L1 or L2, and with or without the help of a mind map. The previously reported L2 cost was replicated, with lower recall scores in L2 relative to L1. Whereas L1 and L2 students were equally aided by AO use, the initial cost dissipated when comparing L2 test scores in the AO condition with those of L1 in the non-AO condition. We therefore conclude that employing AOs does not entirely diminish L2 disadvantages but brings L2 students up to the initial level of L1 students.

Introduction

Within the field of bilingualism, there is ample empirical evidence that second language (L2) processing is less efficient than first language (L1) processing (Cop et al., 2015; Dirix et al., 2020; Vander Beken & Brysbaert, 2018; Whitford & Titone, 2012). For instance, eye-tracking data on bilingual reading patterns reveal longer reading times in L2 compared to L1 at both word and sentence level (Cop et al., 2015; Whitford & Titone, 2012). Subjects in the L2 condition have been found to require 17.6% longer reading times. Furthermore, previous studies into the retention of L2 material have established a significant recall cost for content processed in the L2 (e.g., Vander Beken & Brysbaert, 2018). Firstly, when studying L2 material, subjects take 20% longer to revise the text (Dirix et al., 2020), but importantly, inferior recall outcomes are associated with retention of L2 learning material (e.g., Fecteau, 1999; Vander Beken & Brysbaert, 2018). This L2 cost may possibly be attributed to the processing demands L2 puts on attention and working memory (Francis & Gutiérrez, 2012; see also the resource hypothesis, Sandoval et al., 2010). Within this view, memory performance is modulated by the amount of cognitive resources available, and higher cognitive load results in a smaller capacity for higher order processing. Because L2 processing exploits more resources, it might interfere with other processes that are necessary in the development of strong memory traces. For instance, inference making seems to be strongly reduced during L2 reading (Pérez et al., 2019). Another possibility relates to the frequency with which L2 words are encountered. Within the *weaker links hypothesis* (Gollan et al., 2008), an analogy is drawn to frequency effects that can also be observed in L1. That is, due to their relatively low frequency of occurring to a dominantly L1 speaker, L2 words are processed in very much the same way as low-frequency L1 words. Hence, L2 words will have weaker links with their meanings than words that are more frequently encountered.

Importantly, considering that memory processes can be divided into three subprocesses (i.e., encoding, storage, and retrieval; e.g., Tulving, 1995), it is not yet clear at what stage the L2 cost presents itself. One possibility is that the disadvantage is merely situated at the production or retrieval level. This entails that subjects acquired the necessary knowledge but are unable to reproduce all of it. Another possibility is that the memory cost is situated much earlier in the process, at the encoding stage, and actual differences exist in how much of the L2 material is remembered. If this is the case, a more dramatic picture of L2 content processing emerges. One of the reasons to assume encoding deficits in L2 can be derived from the so-called Landscape Model proposed by van den Broek et al. (1999). The model assumes that mental models constructed from texts consist of landscapes of activated concepts, which in turn activate related concepts to a certain degree. These landscapes are continuously updated during reading. This view is also in agreement with other models assuming that construction of textbase representations is based on background knowledge activation, such as the Construction-Integration model by Kintsch and van Dijk (1978). In accordance with the weaker links hypothesis, one may presume that cohort activation of related concepts will be less profuse, due to fewer and feebler links between concepts. Interestingly, the L2 deficit appears to be greater with respect to central information (i.e., the gist of the text) relative to more peripheral information, suggesting that the use of L2 somehow hampers the construction of a coherent representation of the text (Miller & Keenan, 2011). Furthermore, when Vander Beken and Brysbaert (2018) investigated the effect of test type (recognition vs. recall) on how much information Dutch-English bilinguals correctly remembered from the texts they previously studied in L1 or L2, participants demonstrated similar recognition scores on true/false statements for both language conditions, but lower recall scores on an open questions task in the L2 condition. The fact that no differences were found in true/false judgment may lead to the conclusion that L2 learners are adequately able to encode and store the L2 textual information. Still, an alternative explanation related to differences in the richness or strength of the memory trace cannot be ruled out based on these findings; whereas L2 text comprehension might have been sufficient for recognition, free recall may require a more elaborate memory trace.

In a follow-up study, Vander Beken et al. (2018) addressed the issue of storage by looking at memory for texts in L1 versus L2 over a longer period. They hypothesized that if the mental representation of textual information is indeed weaker in L2 (i.e., differences already emerge at the level of encoding and storage), long-term memory should suffer from additional loss of information, and language differences should also be visible at recognition level. Indeed, Craik and Lockhart (1972) determined that initially weak mental representations are also more prone to disintegration. This implies that L2 readers are at an even greater disadvantage than previously assumed, as the L2 cost may grow even larger over time. Conversely, if the memory cost is due to L2 retrieval deficiencies only, test scores in L1 and L2 should remain similar, even over a period of 1, 7, or 30 days. In support of the latter, memory curves revealed no significant effect of language in the delayed recognition task, which consisted of true/false statements. Yet, as the authors also rightfully acknowledge, the mere fact that the L2 memory trace is as persisting as the L1 memory trace does not directly prove the strength of the representation. Notably, only recognition was measured in this study, whereas the more differentiating variable of recall was not. Evidence suggests that even marginal knowledge—memory content that cannot be retrieved spontaneously but can be recognized or retrieved by cued recall (Berger et al., 1999)—can have very stable memory traces over long periods (see, for instance, Bahrick et al., 1975). Perhaps also weak memory traces can have long-lasting effects.

To ascertain whether L2 memory costs can be located at the richness of the memory trace, the present study's aim was to enhance the mental model for both L1 and L2 material and see how this enhancement affects memory performance. To facilitate meaningful learning and thus promote effective retention, Ausubel (1960) encourages the use of advance organizers (AOs). This involves the use of an organizer prior to presentation of the learning material. By introducing relevant subsuming concepts in advance, organizers provide optimal anchorage and therefore facilitate the incorporation of subsequently presented learning material, which in turn enables (long-term) retention of what is learnt (Luiten et al., 1980). As a coherent mental model of the text is achieved by forming meaningful relationships between different concepts from the text and the reader's prior knowledge (van den Broek et al., 1999), AOs seem particularly suitable to boost the construction of such a solid memory representation. Assuming that organizing tools have a beneficial effect on learners' underlying mental model, the comparison of learners' memory for texts that are learnt either with or without the help of an AO might provide more insight into the origin of the L2 memory cost. If the cost is merely situated at the level of retrieval, L2 learning material is not expected to benefit from an enhanced mental model, as there was no initial problem at this level.

In our attempt to explore the effect of AOs on retention of L1 and L2 learning material, we asked participants to study two text passages in either L1 or L2, with (AO) and without (NAO) the use of organizers. In line with our predictions, several comparisons were of interest: (a) whether memory performance is better in L1 compared with L2 (replication of the results obtained by Vander Beken & Brysbaert, 2018); (b) whether memory performance is better in the AO as opposed to the NAO condition; and (c) whether use of an AO can counteract the L2 recall cost, which would suggest that it may be (partially) due to encoding or storage deficiencies.

Method

Participants

A total of 115 first-year students in psychology at Ghent University with Dutch as L1 and English as L2 participated in exchange for course credit. Fifteen participants were excluded from the analysis, as information gathered during the experiment revealed they did not comply with predefined requirements (i.e., exceeding the age of 30, $N = 4$; not Dutch as L1, $N = 5$; diagnosed with reading/learning disabilities, $N = 6$). The final dataset consisted out of data from 100 participants (age: M = 20.06; SD = 2.01; 67 females), all randomly assigned to a language condition. Informed consent was obtained from all participants prior to the experiment. All subjects were healthy young adults and their participation was entirely voluntarily.

Materials

Learning Material

Participants were presented two scientific texts, "The Sun" and "Sea Otters" (Roediger & Karpicke, 2006). These were translated and adapted by Vander Beken and Brysbaert (2018) to match language variants for semantics and word frequencies. Of each text, there was a Dutch and English variant (word counts between 248 and 279 depending on text variant), printed on paper in Times New Roman 12 with line spacing of 1.5.

Advance Organizers

For all four text variants, a mind map was constructed with the key concept displayed in the center of the page, radiating into several main and subbranches and creating a spider-like web of related concepts (Buzan et al., 2010). The maps provided a schematic overview of the text structure, as well as 20 out of the text's 30 idea units (see Supplementary Material). These 20 ideas were randomly chosen and not more key to the text than the remaining 10. By not including all information in the mind maps, a comparison could be made between recalled ideas that were only in the text and recalled ideas that were both in the text and the mind map, enabling us to see whether participants indeed employed both mind map and text (as they were instructed to), or simply used the map as learning material.

Free Recall Tests

Retention of the two text passages was tested by free recall. Participants were asked to write a full summary of the passages they read.

Distractor Task

As in previous studies, a short delay was implemented between each learning and testing phase by means of a distractor task. Following Vander Beken and Brysbaert (2018), we opted for a computerized version of the Corsi Block tapping task. This visuospatial memory task requires participants to repeat the observed sequence of blocks in the right order (Corsi, 1972). As argued by Vander Beken and Brysbaert (2018), choosing visuospatial stimuli (e.g., block tapping) over verbal stimuli (e.g., multiplication problems) allows us to avoid strong internal activation of a specific language during the task.

L1 and L2 Proficiency

Objective measures of language proficiency were obtained by administering the Lex-TALE lexical decision test (Lemhöfer & Broersma, 2012) in both L1 (Dutch) and L2 (English), in which participants had to indicate whether the presented words were actual words or nonwords. Lemhöfer and Broersma (2012) validated LexTALE by examining its relationship with other measures of L2 proficiency (such as a translation task and a test for general English proficiency; i.e., the Quick Placement Test, or QPT; University of Cambridge, Local Examinations Syndicate, 2001) and found a high correlation between the two. We can therefore conclude that LexTALE can serve as a proxy for general language proficiency, without it being too time-consuming and mentally taxing for the participants. In addition, self-reported proficiency was provided on a 4-point scale (1 = no knowledge, 4 nativelike knowledge) in each of the four skills (comprehension, speaking, reading, writing) for all known languages.

Text-Related Reports

For each text, participants were asked to rate on a 7-point scale (1 = low, 7 = high): (a) prior knowledge of the learning material, (b) experienced text difficulty in terms of content, (c) experienced text difficulty in terms of structure, (d) interest in the text, (e) motivation to do well on the test, (f) degree of AO use, and (g) estimated performance relative to fellow students. At the end of the experiment, they received some additional questions in which they rated their experience with mind maps (1 = low, 7 = high) and their attitude toward English as a medium of instruction (1 = contra, 5 pro). All questions were presented in Dutch.

Design

We opted for a 2 × 2 mixed-factorial design with both within- and between-subjects variables. Learning condition (AO vs. NAO) was manipulated within subjects, the language of the learning material (L1 vs. L2) was manipulated between subjects. The order of the learning conditions, as well as order and language in which learning material ("The Sun" or "Sea Otters") was presented, was counterbalanced across participants. This resulted in eight cells to which participants were randomly assigned.

Learning and Testing Procedure

The experiment was administered in groups of maximum 15 participants. To allow meaningful comparison of results, the learning and testing procedure bore strong resemblance to that of Vander Beken and Brysbaert (2018). At the start of the experiment, participants were informed they would have to study two texts, and their memory of these texts would be tested afterward by means of a free recall test. They were briefed about the fact that one text would be accompanied by a mind map and the other one would not. Hence, they also received some information about the structure of the mind map and an example. It was made clear that the mind map was a learning aid only, and that not all information from the text would be depicted in the map. The test instructions were provided prior to the studying of the text and read as follows: "Write a summary of the text you have just read. Be as detailed as you can be. You do not need to copy the text literally (word for word) but give as much information as you can." For both texts, participants started with a 7-minute learning phase to study the text passage. Participants could see how much time there was left by glancing at a stopwatch displayed in the room. They were allowed to take notes during the entire learning phase. Afterward, they were asked to turn around their sheets and initiate the computerized Corsi block tapping task. After 2 minutes, they were interrupted and went over to the 7-minute testing phase, followed by the text-related questions. This procedure was repeated for the second text. To conclude, participants took the LexTALE for both L1 and L2, and filled out the final part of the questionnaire.

Scoring Procedure

Participants' summaries were scored on recall of the 30 idea units identified by Roediger and Karpicke (2006), resulting in a maximum score of 30 points per text. For each correctly recalled unit, participants received one point. Some of the ideas were subdivided into smaller chunks, allowing fractional scores for partially recalled ideas (e.g., when participants wrote "Sea otters are mustelids" or "Sea otters are the largest of their family" for the idea "Sea otters are the largest of the mustelids," they received a score of 0.5). This subdivision was implemented in response to scoring issues and resulted in a more strictly defined correction key with less room for subjective interpretation. No points were allocated if the idea was missing or incorrectly recalled. The order in which the ideas were reported was not taken into account. Spelling errors and grammatical mistakes were condoned (e.g., allowing for phonetic spelling of a word), as long as they did not result in a change in meaning. All texts were scored by the author, a subset (20%) was independently scored by a second rater.

Results

Interrater and Test Reliability

The reliability of the test rating procedure was evaluated by calculating the intraclass correlation coefficient (ICC) between the first and second rater. We observe an ICC(3) value of .969 (95% CI = [.949, .982], $F(39, 39) = 64.725$, $p < .001$), indicating that use of the correction key results in highly consistent data irrespective of the rater. Cronbach's Alpha was calculated for both recall tests; "The Sun" (L1: $\alpha = .704$; L2: $\alpha = .669$) and "Sea Otters" (L1: $\alpha = .622$; L2: $\alpha = .621$).

Group Comparisons

Language Proficiency

As the experimental design involved a between-group manipulation of language, we first checked whether the two groups were equivalent in terms of L1 and L2 proficiency. As depicted in Table 1, scores on the English version of the LexTALE ($M = 73.66$, $SD =$ 11.30) were substantially lower than those on the Dutch version ($M = 87.83$, $SD = 7.67$). These results are very similar to the proficiency scores reported by Vander Beken and Brysbaert (2018) ($M_{L1} = 89.54$, $M_{L2} = 71.96$) in a comparable group of Dutch-English participants. Importantly, as confirmed with Welch two-sample *t*-tests, there were no significant differences between the two language groups in terms of L1 or L2 proficiency ($p > .05$).

Text-Related Reports

Small differences can be observed between the two language groups in terms of text-related issues (Table 2). Participants in the L2 group rated the learning material as more difficult (both in terms of content and structure) and reported lower levels of prior knowledge, interest, motivation, and self-estimated performance. However, taking into account a correction for multiple comparisons, Wilcoxon rank sum tests indicated that none of these differences were significant ($p > .05$).

Table 1. Group characteristics in absolute values or means (with standard deviations)

	L1 group ($N = 49$)	L2 group ($N = 51$)	Total ($N = 100$)
Gender F/M	32/17	35/16	67/33
Age	20.04 (1.89)	20.02 (2.19)	20.03 (2.04)
LexTALE L1	87.86 (7.83)	87.80 (7.59)	87.83 (7.67)
LexTALE L2	73.70 (11.46)	73.61 (11.26)	73.66 (11.30)

Table 2. Mean ratings on the text-related questions (1 = low, 7 = high), reported per language group and text passage (standard deviations between parentheses)

	L1 group	L2 group	The Sun	Sea otters
Difficulty (content)	3.49 (1.45)	3.76 (1.38)	3.93 (1.42)	3.33 (1.35)
Difficulty (structure)	2.89 (1.36)	3.16 (1.26)	3.02 (1.33)	3.03 (1.31)
Prior knowledge	2.24 (1.48)	1.95 (1.42)	2.55 (1.65)	1.63 (1.06)
Interest	4.88 (1.33)	4.60 (1.50)	4.92 (1.40)	4.55 (1.42)
Motivation	5.47 (1.12)	5.44 (1.30)	5.50 (1.23)	5.41 (1.21)
Estimated performance	4.46 (1.09)	4.11 (0.95)	4.29 (1.05)	4.27 (1.02)

Text Comparisons

Averaging the two language groups, participants indicated more prior knowledge about the Sun than about sea otters ($V = 1617.5$, $p < .001$), but found 'The Sun' more difficult in terms of content ($V = 2153.5$, $p < .005$). No significant differences were observed for structural difficulty, interest, motivation, or self-estimated performance ($p > .05$). Significant differences were also observed in participants' performance on free recall ($t(99) = 7.22$, $p < .001$), with higher test scores for 'The Sun' ($M = 17.01$, $SD = 4.40$) than for 'Sea Otters' ($M = 14.37$, $SD = 3.81$).

Memory Performance

Regarding memory performance, three comparisons were of interest: (a) whether L1 retention was better compared with L2; (b) whether AOs aided retention; and (c) whether AO use counteracts the L2 recall cost. An overview of total test scores as a function of both learning condition and language is provided in Table 3. In line with our prediction, higher test scores were obtained in L1. With respect to the use of AOs, two observations tend to give the impression that this did not affect participants' memory performance. Firstly, the averaged total test scores reveal only small absolute differences between NAO and AO conditions. Secondly, looking at the percentage of participants that performed in line with our prediction (i.e., better test scores with the use of an AO, referred to as Percent Correct Classifications; see Grice et al., 2020), values are close to chance level. This implies that although about half of the participants' behaviour complied with our prediction, the other half of the participants obtained better test scores without the use of an AO.

A more detailed analysis of test scores, however, confirmed that our manipulation did not fail to influence our dependent variable in a consistent manner. As described in the "Method" section, participants' summaries were scored based on the presence and correctness of the 30 ideas from the text. This test score (max = 30) could be subdivided into two subscores; one based on items present in both text and mind map (max = 20) and one based on items present in the text only (max = 10). Closer inspection of these subscores revealed that, despite failure to increase total test score, the use of an AO did affect the *type* of items recalled. That is, relative to the NAO condition, participants

Table 3. Means, medians, standard deviations, ranges of scores (max = 30) on the free recall tests, and percent correct classifications (PCC) as a function of learning condition and language

		Mean	Median	SD	Min	Max	PCC
L1	NAO	16.92	16.85	3.96	6.90	24.50	44.89%
	AO	16.85	17.25	4.52	8.85	25.50	
L2	NAO	14.40	14.00	4.10	4.00	24.00	52.94%
	AO	14.70	15.00	4.13	4.20	22.00	

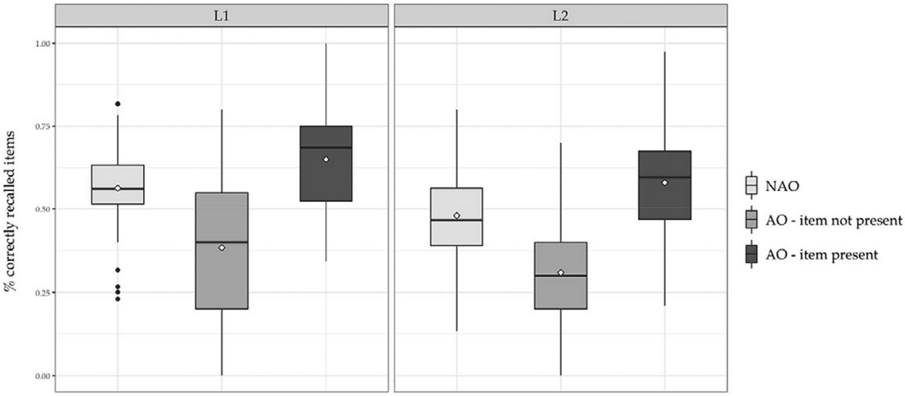

Figure 1. Boxplots of the percentage correctly recalled test items as a function of language (L1 vs. L2) and Item Type (NAO, AO—item not present, AO—item present). Means are represented by the tilted squares.

performed better on the subset of items that were depicted in the mind map (AO: $M =$ 61.47%, $SD =$ 15.87%; NAO: $M =$ 52.12%, $SD =$ 14.02%), but received lower scores for items not depicted in the mind map (AO: M = 34.55%, SD = 22.19%; NAO: M = 49.15%, $SD =$ 20.38%). The same trends can be observed when differentiating between the two language conditions (Figure 1) and are, in contrast to the total test scores, more consistent throughout the sample. For items depicted in the AO, the majority of participants performed better in the AO condition (L1 = 65.31%; L2 = 60.78%); for items not depicted in the AO, the majority of participants performed better in the NAO condition (L1 = 73.47%, L2 = 58.82%).

To account for the heterogeneity of items, formal analysis of memory performance was conducted at item-level (resulting in 30 observations per participant per text). The variable item ranged from 1–60, each value coding for one of the 30 ideas in the two texts. Using the lme4 package (version 1.1-25) in R (R Core Team, 2020), a generalized linear mixed-effect model was fitted with memory performance as binary dependent variable. Items for which at least 50% of the subideas were correctly recalled were allocated a value of 1; items for which less than 50% were correctly recalled were assigned a value of 0.

The variables Language (two levels: L1 vs. L2), AO (two levels: NAO vs. AO), and Item Type (included in the AO vs. not included in the AO) were entered as categorical predictors. To allow for the effect of AO to depend on both Language and Item Type, two- and three-way interaction terms were also included. The initial model contained the maximal random structure justified by the design (cf. Barr et al., 2013). From this model, the smallest random effect was removed to allow for a nonsingular fit. The final random structure contained random intercepts for participants and items, as well as by-participant and by-item random slopes for AO and Item Type. As the model with a three-way interaction was not significantly different from its nested model with only two-way interaction terms ($\chi^2(1) = 1.403$, $p = .236$), the more parsimonious model was retained. Model diagnostics for the final model were evaluated by visual inspection of random effect quantiles and simulated residuals, respectively obtained by use of the R packages sjPlot (Lüdecke, 2018) and DHARMa (Hartig, 2021). There were no signs of distributional violations or overdispersion.

Results for the fixed effects are displayed in Table 4; full output of the final model is provided in the SOM. In line with visual inspection of the data, results indicate a significant interaction between AO and Item Type ($p < .001$). The interaction between

https://doi.org/10.1017/S0272263122000109 Published online by Cambridge University Press

Table 4. Generalized mixed-effects model: Output fixed effects

	Estimate	SE	*z*-value	*p*-value
(intercept)	0.122	0.128	0.949	0.343
Language	0.213	0.067	3.200	**0.001**
AO	0.076	0.046	1.668	0.095
Item Type	–0.497	0.116	–4.286	**<0.001**
AO * Item Type	0.282	0.043	6.535	**<0.001**
AO * Language	–0.002	0.033	–0.044	0.965
Item Type * Language	0.032	0.037	0.860	0.390

Language and AO was not significant ($p = .965$). To explicitly determine whether the use of the AO in L2 aided students to overcome the L2 cost, we compared the L1-NAO condition with the L2-AO condition, for items present in the AO. This posthoc pairwise comparison revealed that the difference between these two conditions was not significant ($p > .05$), indicating that L2 readers did benefit from use of the AO.

Discussion

Previous research into second language (L2) text retention compared to the first language (L1) text retention has determined a recall cost of L2 content material (e.g., Fecteau, 1999; Vander Beken & Brysbaert, 2018). One possible explanation proposes this cost stems from weaker mental representations of textual information due to less efficient processing in L2. The current study aspired more insight into this possible locus and employed the technique of an AO in the form of a mind map to determine whether L2 mental models may be enhanced. To this end, participants were asked to study two short text passages, in either L1 or L2, and with or without the help of the AO. We expected better memory performance for texts written in the learners' L1 (i.e., a replication of the results obtained by Vander Beken & Brysbaert, 2018), as well as a beneficial effect of AO use. We further hypothesized that if use of an AO counteracts the so-called L2 recall cost, its locus may be traced back to encoding issues.

With respect to our first prediction, we were able to replicate the L2 recall cost observed by Vander Beken and Brysbaert (2018). Free recall was indeed better for material presented in L1 compared with L2, denoting a substantial cost of using a nonnative language in educational settings. In addition, and not entirely in accordance with our second prediction, we found that use of a mind map significantly improved performance on the items depicted in the map but was also associated with lower recall scores for other information. As there was no difference in terms of total recall, we are hesitant to draw strong conclusions. Our results show that the AOs were indeed able to boost memory performance, but that this effect was restricted to information that was depicted in the mind map (see also Figure 1). The observation that lower scores were obtained for the text-only subscale could perhaps be explained by the fact that summaries and schematics usually serve as a filter to distinguish important from unimportant information. This might have nudged participants to neglect these items, both during the learning phase (e.g., by putting more effort in the items depicted in the mind map) and the testing phase (e.g., by advertently omitting these items from their summaries as they were considered unimportant). Using AOs in real-life settings, it would indeed be illogical to randomly omit relevant learning material from an instructional map. Following this reasoning, the absence of a beneficial effect of AOs might be due to our somewhat unnatural manipulation, and the positive results for one of our subscales may generalise

https://doi.org/10.1017/S0272263122000109 Published online by Cambridge University Press

to an overall benefit if the mind maps had depicted all test items. However, it could also be that our mind maps did not boost text memory (cf. the nonsignificant difference in total test score), but simply caused a shift in the type of items that were recalled. Even so, the usefulness of mind maps remains evident, as they are able to call attention to specific parts of the text. By doing so, well-designed AOs might be particularly useful to guide students to identify and remember the most important information.

With respect to the origin of the L2 recall cost, we hypothesized that if this cost is due to an L2 deficit in the strength of memory representations, memory-boosting tools should aid participants in their L2 to get to the level of those studying in L1. Although the L2 readers in the AO condition still performed worse than the L1 readers in that same condition, the comparison with L1-NAO showed that L2 readers did benefit from the use of the AO to the point where they performed as well as the L1 readers did without AO. Because it seems unlikely that AOs presented before studying assist learners at the retrieval stage, our findings corroborate the hypothesis that the L2 memory cost may indeed be situated at the level of encoding, and also imply that encoding deficits in L2 can be counteracted. Possibly, the weaker conceptual links (Gollan et al., 2008) were boosted by our memory aid, which brought about cohort activation of related concepts and stimulated the mental landscape (van den Broek et al., 1999). In turn, this may have facilitated the construction of a coherent representation of the text (Miller & Keenan, 2011).

Additionally, the use of mind maps in the current study may also have compensated for encoding difficulties that arose from secondary effects of L2 use, such as diminished interest in the topic or reduced motivation to read and study (see also Duan et al., 2020; Murty & Dickerson, 2017). Indeed, small (but statistically nonsignificant) differences were observed between this study's two language groups in terms of interest, motivation, and self-estimated performance. That is, participants in the L2 group systematically rated these as lower. In addition, they also considered the learning material to be more difficult (both in terms of content and structure) and reported lower levels of prior knowledge. As AOs have been shown to increase learners' motivation for studying (Schaal, 2010), there is the possibility that use of mind maps somewhat counteracted the adverse motivation effect in the L2 condition. However, because the current study employed only one simple question with an ordinal scale to measure motivation, future research should shed more light on this aspect.

Although the outcome of this study shows that L2 users can benefit from tools enhancing mental representations during studying, it also demonstrates that they will never function at the same level as L1 users when given the same means. In addition, further research should determine whether or not this L2 content enhancement has a stable memory trace, lasting over longer periods. Research into the use of mind maps for L1 learning material suggests that this may indeed be the case (Farrand et al., 2002). All in all, given the magnitude of the immediate L2 recall cost, results from the current and previous studies do highlight the need to further explore and address the observed deficits. Care should be taken in choosing the appropriate test format to avoid an underestimation of the learners' knowledge on the topic.

Conclusion

The current study determined a recall cost for L2 learning material when participants were asked to freely recall the content of our texts. Employing AOs in the form of mind maps did not entirely diminish this L2 cost, but did bring L2 students up to the same level as L1 students who were not aided by mind maps. This outcome suggests weaker mental representations during encoding as a possible locus of the L2 disadvantage in immediate text recall.

Notes

Inspection of the excluded data revealed that the participants were randomly spread over the eight cells in our design.

Acknowledgments. This study was made possible by LEMMA (Language, Education, and Memory in Multilingualism and Academia), a Concerted Research Actions fund (GOA—BOF13/GOA/032) of Ghent University, and the Research Foundation—Flanders (FWO Belgium—FWO 1255214N).

Supplementary Materials. To view supplementary material for this article, please visit http://doi.org/10.1017/S0272263122000109.

Ethical Considerations. All subjects were healthy young adults. Their participation was entirely voluntarily and obtained through a document of informed consent. These documents were read and signed before task administration commenced.

Data Availability Statement. The materials and data are available at Open Science Framework: https://osf.io/nh9sb/?view_only=2135b18763d14666bd24da96f682d40e

References

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, *51*, 267–272. https://doi.org/10.1037/h0046669.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology*, *104*, 54–75. https://doi.org/10.1037/0096-3445.104.1.54.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Berger, S. A., Hall, L. K., & Bahrick, H. P. (1999). Stabilizing access to marginal and submarginal knowledge. *Journal of Experimental Psychology: Applied*, *5*, 438–447. https://doi.org/10.1037/1076-898X.5.4.438.

Buzan, T., Buzan, B., & Harrison, J. (2010). *The mind map book: Unlock your creativity, boost your memory, change your life*. Pearson BBC Active.

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PloS One*, *10*, e0134008. https://doi.org/10.1371/journal.pone.0134008.

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International*, *34*, 891B.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Dirix, N., Vander Beken, H., De Bruyne, E., Brysbaert, M., & Duyck, W. (2020). Reading text when studying in a second language: An eye-tracking study. *Reading Research Quarterly*, *55*, 371–397. https://doi.org/10.1002/rrq.277.

Duan, H., Fernández, G., van Dongen, E., & Kohn, N. (2020). The effect of intrinsic and extrinsic motivation on memory formation: Insight from behavioral and imaging study. *Brain Structure and Function*, *225*, 1561–1574. https://doi.org/10.1007/s00429-020-02074-x.

Farrand, P., Hussain, F., & Hennessy, E. (2002). The efficacy of the "mind map" study technique. *Medical Education*, *36*, 426–431. https://doi.org/10.1046/j.1365-2923.2002.01205.x.

Fecteau, M. L. (1999). First-and second-language reading comprehension of literary texts. *The Modern Language Journal*, *83*, 475–493. https://doi.org/10.1111/0026-7902.00036.

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*, 496–503. https://doi.org/10.3758/s13421-011-0163-3.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*, 787–814. https://doi.org/10.1016/j.jml.2007.07.001.

Grice, J. W., Medellin, E., Jones, I., Horvath, S., McDaniel, H., O'lansen, C., & Baker, M. (2020). Persons as effect sizes. *Advances in Methods and Practices in Psychological Science*, *3*, 443–455. https://doi.org/10.1177/2515245920922982.

Hartig, F. (2021). DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.4.1. https://CRAN.R-project.org/package=DHARMa

Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review*, *85*, 363–394. https://doi.org/10.1037/0033-295X.85.5.363.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*, 325–343. https://doi.org/10.3758/s13428-011-0146-0.

Lüdecke, D. (2018). sjPlot: Data visualization for statistics in social science. R package version 2.8.7. https://CRAN.R-project.org/package=sjPlot

Luiten, J., Ames, W., & Ackerson, G. (1980). A meta-analysis of the effects of advance organizers on learning and retention. *American Educational Research Journal*, *17*, 211–218. https://doi.org/10.3102/00028312017002211.

Miller, A. C., & Keenan, J. M. (2011). Understanding the centrality deficit: Insight from foreign language learners. *Memory & Cognition*, *39*, 873–883. https://doi.org/10.3758/s13421-010-0062-z.

Murty, V. P., & Dickerson, K. C. (2017). Motivational influences on memory. In *Recent Developments in Neuroscience Research on Human Motivation Advances in Motivation and Achievement* (Vol. *19*, 203–227). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-742320160000019019.

Pérez, A., Hansen, L., & Bajo, T. (2019). The nature of first and second language processing: The role of cognitive control and L2 proficiency during text-level comprehension. *Bilingualism: Language and Cognition*, *22*, 930–948. https://doi.org/10.1017/s1366728918000846.

R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*, 231–252. https://doi.org/10.1017/S1366728909990514

Schaal, S. (2010). Enriching traditional biology lectures-Digital concept maps and their influence on achievement and motivation. *World Journal on Educational Technology*, *2*, 42–54.

Tulving, E. (1995). Organization of memory: Quo vadis? In M. S. Gazzaniga, *The Cognitive Neurosciences* (pp. 753–847). MIT Press.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Erlbaum.

Vander Beken, H., & Brysbaert, M. (2018). Studying in a second language: The importance of test type. *Bilingualism: Language & Cognition*, *21*, 1062–1074. https://doi.org/10.1017/S1366728917000189.

Vander Beken, H., Woumans, E., & Brysbaert, M. (2018). Studying texts in a second language: No disadvantage in long-term recognition memory. *Bilingualism: Language & Cognition*, *21*, 826–836. https://doi.org/10.1017/S1366728917000360.

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review*, *19*, 73–80. https://doi.org/10.3758/s13423-011-0179-5

Cite this article: Dhaene, S. and Woumans, E. (2023). Text recall and use of advance organisers in first and second language. *Studies in Second Language Acquisition,* **45,** 264–275. https://doi.org/10.1017/S0272263122000109