tesol QUARTERLY

The Impact of Gaze-Contingent Textual Enhancement on L2 Collocation Learning from Computer-Mediated Reading Tasks

JOOKYOUNG JUNG
The Chinese University of Hong Kong
Sha Tin, Hong Kong SAR

ANDREA RÉVÉSZ
University College London
London, UK

MATTHEW J. STAINER
Griffith University
Gold Coast, Australia

ANA PELLICER-SÁNCHEZ
University College London
London, UK

YOOJIN CHUNG
University College London
London, UK; and
University of Reading
Reading, UK

DANNI SHI
Georgetown University
Washington, USA

Abstract

This study examined if gaze-contingent textual enhancement could be used as an interactive focus-on-form device to promote learning of second language (L2) collocations from computer-mediated reading tasks. Seventy-five Chinese ESL users read three English texts that contained twelve target collocations, presented under one of three conditions: no highlighting, proactive highlighting (target collocations highlighted in advance), and gaze-contingent highlighting (target collocations highlighted when looked at). Participants' eye movements were captured during the reading task, and collocation form recall and recognition tests were administered immediately after and 2 weeks later. Additionally, five participants from each group took part in a stimulated recall session, eliciting their thoughts while reading. The results indicated that both highlighting techniques increased total fixation duration and count on the target collocations and improved collocation form recall and recognition scores in the posttests. Gaze-contingent highlighting demonstrated a more durable impact on the collocation recall test compared to proactive highlighting. The stimulated recall comments also revealed that gaze-contingent highlighting tended to promote attentive processing of the target collocations. These findings suggest that highlighting is a useful focus-on-form technique in task-based reading contexts, with gaze-contingent highlighting yielding potential benefits in terms of L2 collocation learning.

doi: 10.1002/tesq.3404

INTRODUCTION

Eye-tracking technology has gained considerable attention from second language (L2) researchers as a tool for investigating learners' attentional processing of visual information displayed on computer screens (Conklin, Pellicer-Sanchez, & Carrol, 2018; Godfroid, 2019). Due to its high temporal and spatial resolution, eye-tracking has widely been used to capture learners' real-time visual processing (Roberts & Siyanova-Chanturia, 2013; Winke, 2013). Recently, eye-tracking technology has also begun to be used as an attention trigger (Révész, Stainer, Jung, Lee, & Michel, 2023), not only as an attention tracker, based on the gaze-contingency paradigm (Reder, 1973). Within this paradigm, the visual stimulus presented on the computer screen dynamically reacts to the viewer's gaze (or lack of it), enabling the creation of learner-adaptive programs that direct attention toward specific on-screen elements, such as targeted images or texts (Duchowski, 2003). Révész et al. (2023) highlighted that the integration of learner-adaptive programs utilizing the gaze-contingent paradigm aligns well with the focus-on-form approach (Long & Robinson, 1998), making it possible to direct learners' attention to target linguistic features during communicative tasks in response to learners' needs. Indeed, gaze-contingent glossing, a type of focus-on-form technique, has been found to facilitate attention to and learning of single L2 words (Révész et al., 2023). It remains unexplored, however, whether these observed benefits extend to other focus-on-form interventions and linguistic features.

To examine this possibility, this study aimed to investigate the effectiveness of gaze-contingent textual enhancement in facilitating the learning of L2 collocations during computer-mediated reading. Textual enhancement, such as underlining, boldfacing, or highlighting, has been widely used as a focus-on-form technique to draw learners' attention to target features during reading (Leow, 2015). However, most existing textual enhancement techniques have been employed proactively, with target features being manipulated before learners' reading. Given the growing interest in the gaze-contingent paradigm (Révész et al., 2023), it appears timely to explore the efficacy of interactive textual enhancement, where enhancement occurs reactively in response to learners' eye gaze. In this study, to achieve a comprehensive understanding of the effects of gaze-contingent textual enhancement, we triangulated eye-movement data, stimulated recall comments, and collocation recall and recognition scores to generate practical guidance for designing learner-adaptive L2 tasks in computer-mediated learning contexts.

LITERATURE REVIEW

Gaze-Contingency Paradigm

Within the gaze-contingent paradigm (Reder, 1973), learners' eye gaze serves as a trigger for visual events on the computer screen, facilitating enhanced accuracy and adaptability during learner-treatment interactions (Wilms et al., 2010). Thanks to this real-time reciprocity of gaze-dependent interface (Duchowski, 2003), this technology has been employed in studies that explored interpersonal abilities that require ocular processing of visual stimuli such as facial cues (e.g., Wang et al., 2020), social signals (e.g., Vernetti et al., 2020), and joint attention (Little, Bonnar, Kelly, Lohan, & Rajendran, 2016). For example, Wang et al. (2020) demonstrated that a gaze-contingent cueing program can help young children with autism spectrum disorder (ASD) to better focus on human faces and hence develop socially acceptable attention-paying skills. Similar findings have been reported in recent studies on autism (e.g., Evers, Van Belle, Steyaert, Noens, & Wagemans, 2017) and infant cognitive and social development (e.g., Keemink, Jenner, Prunty, Wood, & Kelly, 2020), suggesting that the gaze-contingent paradigm can be applied to various forms of learning.

In the context of L2 learning, Lee, Kanakogi, and Hiraki (2015) explored if a gaze-interactive system would help L1 Japanese adult speakers learn L2 Korean words. In their gaze-contingent learning program, participants were guided to make eye contact with an animated three-dimensional cartoon character. This character first made eye contact with the learners and then turned its gaze to the target picture. When the learners looked at the picture and hence formed joint attention, the character returned its gaze to the learner and spoke a framed sentence that contained the target Korean word, followed by a two-time repetition of the word. The results revealed that the gaze-contingent group scored significantly higher on the post-treatment test than those who viewed pre-recorded instructional videos. The findings of this study demonstrate that attentional control using a gaze-contingent learning program has the potential to be utilized as a language instructional tool.

The study closest in focus to the present research was conducted by Révész et al. (2023). This study was the first to examine whether gaze-contingent highlighting could enhance the learning of lexical items from computer-mediated reading. In this study, glosses were highlighted when participants fixated on the target words but skipped visiting the corresponding glosses. This way, gaze-contingent highlighting was used as an attention-triggering device, encouraging learners to notice the glosses and thereby connect the target word forms with their meanings. For the comparison group, participants read texts with traditional non-interactive glosses. The results revealed that the gaze-contingent highlighting could promote visual attention to the glosses significantly, as manifested in significantly higher fixation counts and longer durations than the non-interactive glossing condition. However, gaze-contingent highlighting did not result in greater gains in terms of form and meaning recognition. Notably, more fixations at the glosses were related to better form recognition scores under the gaze-contingent condition but were associated with lower form recognition scores in the non-interactive, non-gaze-contingent condition. Clearly, further research is needed to explore the pedagogical potential of gaze-contingent textual enhancement in promoting L2 learning from computer-mediated reading.

Textual Enhancement and Collocation Learning

Collocations can be defined as word combinations that frequently occur together (Nesselhauf, 2005), and thus collocation knowledge includes the ability to recognize and predict which words typically co-occur (Nation, 2013). To develop collocational competence, learners need exposure to large amounts of input in which they can encounter collocations sufficiently frequently and thereby establish a lexical priming mechanism (Northbrook & Conklin, 2019). For many L2 learners, however, it is difficult to have frequent exposure to collocations, and hence improving collocational knowledge often poses a challenge in L2 learning (Boers, Demecheleer, Coxhead, & Webb, 2014). In addition, even when the component words that make up a collocation are known to learners, they may fail to notice the phraseological connection among the words (Carrol & Conklin, 2014). To address this limitation, textual enhancement has been explored as a focus-on-form technique to promote learners' attention to target collocations in meaning-focused reading tasks (Webb, Newton, & Chang, 2013).

Thus far, diverse enhancement techniques have been employed by researchers, including boldfacing (e.g., Choi, 2017; Toomer & Elgort, 2019), underlining (e.g., Puimège, Montero Perez, & Peters, 2021; Szudarski & Carter, 2016), and coloring (e.g., Jung, Stainer, & Tran, 2022); and sometimes, enhancement techniques have been combined, such as boldfacing with underlining (e.g., Majuddin, Siyanova-Chanturia, & Boers, 2021; Peters, 2012) or with coloring (e.g., Sonbul & Schmitt, 2013). Overall, previous studies have shown that textual enhancement can promote noticing (Choi, 2017; Jung et al., 2022; Majuddin et al., 2021; Puimège et al., 2021; Puimège, Montero Perez, & Peters, 2024) and learning of enhanced collocations (Sonbul & Schmitt, 2013; Szudarski & Carter, 2016; Toomer & Elgort, 2019). For example, Choi (2017) examined whether textual enhancement through boldfacing would drive learners' attention to collocations by measuring their eye movements during reading. The results indicated that boldfaced collocations received significantly more attention from learners than unenhanced ones, which further led to higher learning scores in the posttests. However, the enhancement group recalled significantly less text content than the unenhanced group, suggesting that textual enhancement needs to be conducted carefully, not to induce a trade-off between attention to target collocations and comprehension of the reading material.

A few studies further explored whether textual enhancement would have differential effects on developing implicit and explicit knowledge of target collocations (Sonbul & Schmitt, 2013; Toomer & Elgort, 2019). For example, Sonbul and Schmitt (2013) compared the relative efficacy of decontextualized exposure, input flood, and textual enhancement (red and boldfaced). Explicit learning was assessed using collocation form recall and recognition tests, while implicit knowledge was measured with a primed lexical decision task. The results showed that textual enhancement was more effective than input flood in promoting explicit knowledge of target collocations. Toomer and Elgort (2019) replicated Sonbul and Schmitt's (2013) study while including three input conditions: unenhanced, bolding, and bolding-plus-glossing. The results revealed that textual enhancement promoted explicit knowledge of the target collocations, while the development in implicit knowledge was observed in the unenhanced condition.

The brief overview of studies demonstrates that textual enhancement, functioning as an attention-triggering device, can aid L2 learners in noticing and learning target collocations from reading. It seems also worth noting that previous studies employed proactive textual enhancement, where target features were modified typographically before reading. In contrast, little is known about the impact of textual enhancement triggered in reaction to learners' eye fixations based on the gaze-contingent paradigm (Révész et al., 2023). Gaze-contingent textual enhancement occurs precisely when learners' eye gaze reaches a particular point during reading, ensuring temporal synchronization between their gaze and the timing of textual enhancement. This real-time reciprocity of gaze-contingent textual enhancement may yield distinctive outcomes in capturing learners' attention and promoting L2 collocation learning.

THE CURRENT STUDY

Against this background, the present study examined the effects of gaze-contingent highlighting, in comparison to proactive highlighting, on learners' attentional processes while completing computer-mediated reading tasks and on their learning of L2 collocations embedded in the reading texts. In this study, proactive highlighting was operationalized as highlighting target L2 collocations in advance of reading; whereas interactive, gaze-contingent highlighting was operationalized as highlighting target L2 collocations when participants' eye fixations were detected on them during reading (Révész et al., 2023). More specifically, the aim of this project was to address the following research questions:

1. To what extent do proactive and interactive highlighting affect L2 learners' attention to target collocations embedded in reading texts?
2. To what extent do proactive and interactive highlighting affect L2 learners' development in the knowledge of the target collocation forms?

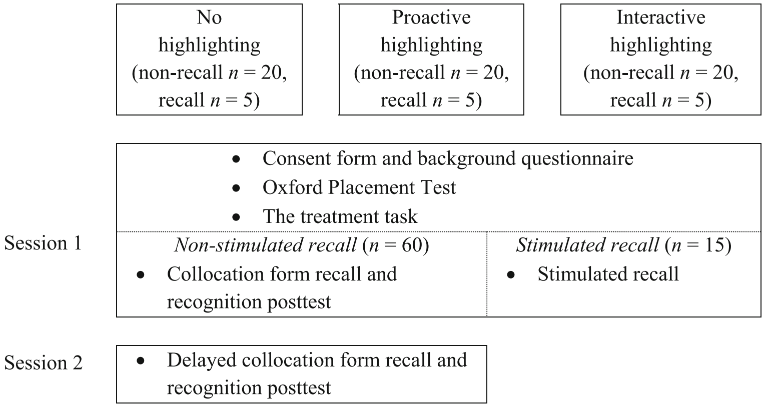

FIGURE 1. Overall design of this study.

METHODOLOGY

Overall Design

The present study adopted a treatment-immediate posttest-delayed posttest design, with participants randomly assigned to three groups (see Figure 1). The between-subjects variable was the textual modification technique (proactive highlighting, interactive highlighting, or no highlighting). The dependent variables included participants' attention to the target collocations operationalized through eye-tracking measurements and participants' knowledge of the target collocation forms assessed through collocation recall and recognition tests. That is, this study focused on changes in participants' ability to identify and acquire the phraseological bond between constituent words within collocations (Nation, 2013). Additionally, five students from each group (n = 15) were invited to engage in a stimulated recall interview immediately after the reading task, instead of taking the posttests.

Participants

The participants were 75 university students in the United Kingdom, who were L1 Chinese and L2 English speakers. We originally recruited 77 participants, but needed to exclude two due to severe drift issues in the eye-movement data. The participants were studying toward various degrees, such as statistics, medicine, education, engineering, and digital media. Considering the topic of the reading materials (economic trends), students with economics majors were excluded from the study. We recruited participants with scores ranging from 6.5 to 7.5 or TOEFL scores between 94 and 109 to ensure that they can comprehend the treatment texts without difficulty. To obtain a more accurate English proficiency measure, the participants took the language use section of the Oxford Placement Test (Cronbach's alpha = .856). They on average scored 75.97 with a maximum score of 100 ($SD = 9.40$), with the large majority scoring in the B2 CEFR range. A one-way ANOVA further revealed that there was no significant difference among the three groups in terms of English proficiency, $F(2, 58) = 1.533$, $p = .225$, $\eta^2 = .053$.

Reading Materials and Target Collocations

The reading materials for this study were three articles on economic trends in 2022, such as inflation combined with supply disruption, blockchain money and art, and increased momentum for the eco-friendly movement. The average length of the three articles was approximately 765 words, and 96.7% of the words were within the most frequent 3,000 words according to the Corpus of Contemporary American English (COCA), and 93.32% within the B2 CEFR level according to Text Inspector (2023). The articles were piloted with six L1 Chinese speakers with comparable English proficiency levels and English learning profiles, and the results revealed that the lexical coverage of the texts was appropriate for the participants.

In this study, a collocation was operationalized as a conventionalized lexical chunk that frequently appears in economic contexts. Each article contained four economic collocations that were relevant to the topic of the article (see Table 1), and each of the target collocations appeared three times in the texts. The 12 target collocations were selected based on the results of a pilot study. In the pilot, 32 L1 Chinese speakers, who had similar English learning profiles to the participants of the current study, received a list of 20 candidate collocations from the same topic area (e.g., *baby bust, debt trap*…). They were asked to mark if the collocations were familiar to them on a 7-point Likert scale and write down the meaning of the collocations if they could. We also confirmed that the constituent words included in the 20 collocations were known to all of the pilot participants. Those collocations that were rated as familiar (above 3) or whose correct meaning was provided by any of the participants were removed from the list. As a result, 12 collocations were selected as the target collocations for the present study. We ensured that each target collocation was accompanied by explicit and transparent contextual information. We

TABLE 1
Target Collocations

Articles	Target collocation	Meaning
Inflation and supply disruption	Bear trap	False signal of a reversal from a down- to an up-market
	Retail cooling	Fall in the value of sales
	Dark store	Warehouse specialized in online shopping delivery
	Micro fulfillment	Small-scale warehouse in urban areas closer to the consumers
Blockchain money and art	Cold wallet	Hardware device for storing coin money
	Diamond hands	Investor who refrains from selling despite loss
	Generative art	Art in which the artist shares the process of art generation
	Right clicker	Those who use the mouse right click to save images
Eco-friendly movement	Vertical farm	Growing crops in vertically stacked layers
	Blue carbon	Carbon stored in coastal and marine ecosystems
	Water footprint	The amount of water you consume in your daily life
	Range anxiety	The fear of running out of battery power on a journey

confirmed this through piloting and made any necessary adjustments when necessary.

Reading Task

Participants were asked to assume the role of an editor for a lifestyle magazine, evaluating the three articles for publication in the next issue. In preparation for this, they received a feedback form before reading so that they could bear the evaluation criteria in mind while reading. They were also allowed to switch freely between the feedback form and the articles. This task-based reading activity aimed to increase the ecological validity and pedagogical relevance of this study by closely aligning the treatment task with real-world tasks. Based on piloting with six L1 Chinese speakers, 15 minutes were assigned for evaluating each article. After participants submitted the feedback form, they completed 10 true-or-false statements about the article that they had just read. This activity was framed as a quality check by the publishing company, assessing whether the editors paid close attention to the content of the articles (Cronbach's alpha = .664). When answering the true-or-false items, participants were not allowed to refer to the article to avoid further exposure to the target collocations.

Highlighting and Eye-Tracking Procedure

In the eye-tracking experiment, the articles were presented in landscape mode, using Courier 14-size font, double-spacing, and left alignment. To assist participants in navigating the texts, page numbers were specified at the bottom of each page (e.g., 1/5, 2/5). We made sure that the two components of each target collocation appeared on the same line, not at the beginning or at the end of a sentence, or next to a punctuation mark (Sagarra & Seibert Hanson, 2011). For the proactive condition, the target collocations were highlighted in yellow before reading. Under the gaze-contingent condition, however, highlighting was only triggered in response to participants' eye fixations. Participants' eye fixations during reading were captured with an EyeLink 1000 Plus eye tracker, recording monocularly at 2,000 Hz. Participants were positioned in front of a 21-inch monitor, maintaining a viewing distance of around 60 cm. They were asked to place their face on a chin rest, which was to stabilize their head movement for obtaining reliable eye-movement data. Before each eye-tracking session, a nine-point calibration and validation procedure was performed. Participants were given sufficient break time between articles, and following each break, the eye tracker was recalibrated.

The Assessment Tools

This study employed paper-based collocation form recall and recognition tests to measure development in the knowledge of the target collocation forms. The collocation recall test was a cloze task in which each target collocation was embedded within a sentence that was taken from the article. The second word was given, and the participants were asked to recall the first component of the collocation. At the piloting stage, L1 Chinese users of English evaluated if the sentences provided sufficient contextual information; the results revealed that they had no difficulty in identifying the meaning of the target collocations. We decided not to give the number of letters for each blank, as the pilot study suggested that this would substantially assist participants. Figure 2 shows a sample recall item for the collocation *blue carbon*.

The collocation form recognition test consisted of multiple-choice items. The same sentences used in the recall test were provided to the participants. The possible responses included the correct first word of

```
Scientists proved that ( _____ carbon) in oceanic areas
can make climate change less severe.
```

FIGURE 2. Sample collocation form recall test item.

```
Scientists proved that ( _____ carbon) in oceanic areas can
make climate change less severe.

a. pure
b. blue
c. clean
d. green
e. tide
f. I DON'T KNOW
```

FIGURE 3. Sample collocation form recognition test item.

the target collocation and four semantically plausible distractor words that could make sense within the sentence. Each item further included an "I don't know" option, which was to discourage random guessing. Figure 3 displays a sample recognition item for *blue carbon*.

The same items were used in the delayed posttest, but to minimize potential practice effects, the order of the items was scrambled. Each of the collocation recall and recognition tests contained 12 items in total, and participants received 1 point for each correct response (Cronbach's alpha: Recall = .901, Recognition = .847).

Stimulated Recall Protocol

As eye-movement data provide limited insights into the depth or the type of cognitive processes associated with eye-gaze behaviors (Godfroid, Boers, & Housen, 2013; Winke, 2013), we triangulated the eye-movement data with stimulated recall comments (Jung & Lee, 2022; Jung & Révész, 2018; Wang & Pellicer-Sánchez, 2023). Once the computer-mediated reading session was completed, five participants from each group ($n = 15$) were asked to attend a stimulated recall session. During the interview, we asked them to recall their thought processes during reading, prompted by the recordings of their eye movements. Participants were encouraged to stop the recording whenever they remembered what they were thinking at the time of reading. The researchers also stopped the recordings if any notable eye movements were observed, such as longer fixations or regressive eye movements, but were not commented on by the participants. The interviews were conducted in Chinese to allow the participants to report their thought processes freely without a linguistic barrier.

Procedure

As visualized in Figure 1, there were two sessions with a 2-week interval. In Session 1, 75 participants signed the consent form and completed a background questionnaire, followed by the Oxford Placement Test. Once randomly assigned to one of the three groups, they performed the editor tasks. Next, those who were assigned to the stimulated recall session were asked to recall what they were thinking while performing the tasks, prompted by their eye movements. The rest of the participants completed a collocation form recall and recognition test immediately and two weeks later. All sessions were conducted individually at a computer lab. Session 1 took approximately 90 minutes and Session 2, 45 minutes. Participants also took working memory and language aptitude tests, which are the focus of a different article.

ANALYSIS

Eye-Movement Measurements

After a drift check, two participants were removed from the data set due to the gaze-contingent highlighting malfunctioning more than 10% of the time. Next, to answer the first research question, interest areas (IAs) were defined for each target collocation, and the total fixation durations (i.e., the combined duration of all fixations) and the total fixation counts (i.e., the overall number of all fixations) captured for each IA were extracted using the SR Data Viewer software (SR Research, www.sr-research.com). These measures were expected to reflect participants' noticing of the target collocations during the word-to-text integration process (Pellicer-Sánchez, Siyanova-Chanturia, & Parente, 2022). Prior to analysis, fixation data shorter than 50 ms (2.05% of the total data) were regarded as outliers and removed from the data set (Conklin et al., 2018; Godfroid, 2019).

Statistical Analysis

The statistical analyses were conducted by constructing mixed-effects models using the package *lme4* (Bates et al., 2014) with the research software R (R Core Team, 2022). The null models had only the random effects, subject and item. For the first research question, we constructed linear mixed-effects models using the function *lmer*, including group as the fixed effect and total fixation durations and counts as the dependent variables, and subject and Item as the random effects. For the second research question, we utilized a series of generalized logistic mixed-effects models using the function *glmer* together with the argument *family* = *binomial* (Linck & Cunnings, 2015). The dependent variables were the binary scores obtained on the collocation form recall and recognition tests. For each mixed-effects analysis, the maximal models were constructed by including the fixed effects as random slopes for all the random effects (Barr, Levy, Scheepers, & Tily, 2013). If the maximal model failed to converge due to an overly complex internal structure, random effect parameters were removed stepwise until convergence was reached (Cunnings & Sturt, 2014). For each model, assumption checks were conducted by examining the normality of the model residuals (for results, see Appendix S1). An alpha level of $p < .05$ was adopted as the benchmark for significance. Effect size estimates for the linear models were obtained using the function *r.squaredGLMM* of the package *MuMln* (Barton, 2015), with R^2 values of .20, .33, and .50 considered small, medium, and large (Plonsky & Ghanbar, 2018). The R^2 value indicated the proportion of variance in the dependent variable explained by the independent variables. For the generalized logistic models, effect sizes were expressed as odds ratios using the package *oddsratio* (Schratz, 2021). An odds ratio greater than 1 indicates that the event is more likely to occur in the experimental group than the control group, while an odds ratio less than 1 indicates that the event is less likely to occur in the experimental group.

Stimulated Recall Comments

The stimulated recall comments were transcribed verbatim and coded using MAXQDA through bottom-up annotations by the first author. In total, eight meta-codes emerged: *task management*, *highlighting*, *word/ phrase*, *pausing*, *slow reading*, *fast reading*, *re-reading*, and *no reason/memory*. *Task management* included participants' comments on perceived task demands, such as whether they could concentrate on the task, whether they were confident about their task performance, or whether they felt the task was difficult. *Highlighting* involved comments related to noticing the proactive or interactive highlighting while reading. *Word/ phrase* referred to comments on noticing a new word or a phrase, inferring its meaning, and evaluating the inferred meaning. When it comes to codes related to reading behaviors, *pausing* included comments that explained why participants paused during the task, and *re-reading* denoted comments about reading the same part of the text again. Also, s*low reading* and *fast reading* involved comments related to the speed of reading. Finally, *no reason/memory* was annotated when participants did not recall their reading behaviors or the reasons for them. A subset of the transcripts (13.33%) was double-coded by an experienced Chinese-speaking applied linguist, and the agreement between the coders was at Cohen's kappa = .885, $p < .001$.

RESULTS

Preliminary Analyses

To determine whether highlighting had any trade-off effects, such as distracting participants' attention from the content of the articles, the impact of highlighting on the true-or-false comprehension items was examined. The results showed that each group achieved reasonably good scores on the comprehension test, indicating that the participants read the articles with a focus on meaning (see Appendix S2 for descriptive statistics). One-way ANOVAs confirmed that the three groups achieved similar scores, Text 1: $F(2, 73) = .1.212$, $p = .305$, $\eta^2 = .042$; Text 2: $F(2, 73) = 1.533$, $p = .225$, $\eta^2 = .053$; Text 3: $F(2, 73) = 1.241$, $p = .297$, $\eta^2 = .043$. These results suggest that neither proactive nor interactive highlighting disrupted participants' processing of the articles for meaning.

Impact of proactive and interactive highlighting on attention to target collocations. Table 2 displays descriptive statistics for the total fixation duration and the total fixation count. To examine the effects of proactive or interactive highlighting on participants' visual attention to the target collocations, we first constructed maximal mixed-effects models for fixation duration and fixation count using the no highlighting condition as the reference. Next, we reran the models using the proactive highlighting group as the reference.

As summarized in Table 3, the results revealed that both proactive and interactive highlighting led to a significant increase in total fixation duration ($R^2_m = .069$, $R^2_c = .363$) and total fixation count ($R^2_m = .065$, $R^2_c = .383$) in comparison to the no highlighting condition. When the two highlighting conditions were compared, however, no significant differences were found. In other words, highlighting the target collocations, either proactively or interactively, helped promote visual attention to the target collocations, with the two highlighting techniques demonstrating similar effects.

TABLE 2

Descriptive Statistics for Eye Movement Measurements Per Target Collocation

Index	Group	*n*	Mean	*SD*	95% CI
Total duration (ms)	No highlighting	20	3598.33	2596.39	[2460.43, 4736.23]
	Proactive highlighting	20	3684.47	2481.08	[2597.11, 4771.83]
	Interactive highlighting	20	4415.58	2508.12	[3316.37, 5514.79]
Total count	No highlighting	20	13.68	8.74	[9.85, 17.51]
	Proactive highlighting	20	15.15	9.72	[10.89, 19.41]
	Interactive highlighting	20	17.13	9.44	[12.99, 21.27]

TABLE 3

Summary of the Best-Fit Models for Eye Movement Measurements

		Fixed effects				Random effects		
		B	*SE*	*t*	*p*	By subject *SD*	By text *SD*	By IA *SD*
Reference group: No highlighting								
Total Duration	Intercept	2907.1	385.5	7.542**	.001	924.8	431.5	628.0
	Proactive	915.8	395.5	2.316*	.027	2022.1	–	–
	Interactive	1505.4	411.7	3.657***	<.001	1369.2	–	–
Total fixation	Intercept	11.821	1.614	7.323**	.006	3.224	2.006	2.682
Count	Proactive	3.866	1.473	2.625*	.014	4.905	–	–
	Interactive	5.551	1.658	3.347**	.002	5.609	–	–
Reference group: Proactive highlighting								
Total Duration	Intercept	3822.9	444.7	7.024***	<.001	1322.9	431.7	628.0
	Interactive	589.6	467.5	1.261	.216	1883.4	–	–
	No highlight	−915.8	395.5	−2.316*	.027	1991.4	–	–
Total fixation	Intercept	15.688	1.856	8.452***	<.001	5.066	2.006	2.682
Count	Interactive	1.685	1.894	.889	.380	.978	–	–
	No highlight	−3.866	1.473	−2.625*	.014	8.288	–	–

Note. Significance level: $*p < .05$, $**p < .01$, $***p < .001$.

Impact of proactive and interactive highlighting on knowledge of the target collocation forms. As shown in Table 4, the descriptive statistics for the immediate and delayed posttests revealed that participants generally achieved higher scores on the recognition tests compared to the recall tests. Additionally, there was a tendency for scores to decrease in the delayed posttests.

To examine the impact of highlighting on participants' knowledge of the target collocations, we constructed mixed-effects models for the

TABLE 4
Descriptive Statistics for Target Collocation Posttest Scores

			Recall			Recognition		
Condition	Session	*n*	Mean	*SD*	95% CI	Mean	*SD*	95% CI
No	Immediate	20	3.05	2.39	[2.00, 4.10]	6.55	2.63	[5.40, 7.70]
Highlighting	Delayed	20	2.75	2.65	[1.59, 3.91]	5.80	2.71	[4.61, 6.99]
Proactive	Immediate	20	4.41	2.45	[3.34, 5.48]	8.76	2.11	[7.84, 9.68]
Highlighting	Delayed	20	3.29	2.82	[2.05, 4.53]	7.59	1.91	[6.75, 8.43]
Interactive	Immediate	20	5.76	3.19	[4.36, 7.16]	8.14	3.02	[6.82, 9.46]
Highlighting	Delayed	20	4.43	3.28	[2.99, 5.87]	7.76	2.62	[6.61, 8.91]

Note. Maximum score = 12.0.

immediate and delayed recognition and recall scores, first using the no highlighting condition and then the proactive highlighting condition as the reference group. As summarized in Table 5, both proactive and interactive highlighting significantly improved immediate collocation recall scores (proactive: odds ratio = 2.671 with 95% CI [1.048, 6.811]; interactive: odds ratio = 4.619 with 95% CI [1.676, 12.733]), with a larger effect size for the comparison between the no highlighting and interactive condition. Also, only the interactive group retained a significant advantage over the no highlighting group on the delayed collocation recall scores (interactive: odds ratio = .463 with 95% CI [.132, 1.621]). Notably, however, the direct comparisons between the two highlighting conditions did not yield a significant difference.

Turning to the collocation recognition scores, both the proactive and interactive highlighting groups outperformed the no highlighting group on the immediate posttest (proactive: odds ratio = 3.181 with 95% CI [1.655, 6.115]; interactive: odds ratio = 2.795 with 95% CI [1.417, 5.515]) and the delayed posttest (proactive: odds ratio = .844 with 95% CI [.448, 1.590]; interactive: odds ratio = .542 with 95% CI [.268, 1.095]). Similar to the recall scores, no significant difference in collocation recognition scores emerged between the two highlighting conditions when directly compared.

In sum, both proactive and interactive highlighting, as compared to no highlighting, assisted participants in achieving receptive knowledge of the target collocations. The two highlighting conditions also facilitated superior gains in productive collocational knowledge over the no highlighting condition in the shorter term. However, only interactive highlighting led to a longer-term advantage over no highlighting at the productive level.

TABLE 5
Summary of the Best-Fit Models for Posttest Scores

		Fixed effects				Random effects	
		B	*SE*	*z*	*p*	By subject *SD*	By item *SD*
Reference group: No highlighting							
Recall Immediate	Intercept	−1.658	.501	−3.307***	<.001	1.302	1.213
	Proactive	.983	.478	2.057*	.040	.819	–
	Interactive	1.530	.517	2.958**	.003	.967	–
Recall Delayed	Intercept	−2.001	.552	−3.625***	<.001	1.622	1.176
	Proactive	.380	.614	.618	.536	1.296	–
	Interactive	1.275	.596	2.138*	.033	1.321	–
Recognition Immediate	Intercept	.043	.331	.129	.897	.595	.922
	Proactive	1.157	.333	3.471***	<.001	.859	–
	Interactive	1.028	.347	2.965**	.003	.834	–
Recognition Delayed	Intercept	−.085	.359	−.237	.813	.933	.870
	Proactive	.721	.335	2.151*	.031	.627	–
	Interactive	.939	.371	2.531*	.011	.712	–
Reference groupo: Proactive highlighting							
Recall Immediate	Intercept	−.676	.474	−1.426	.154	1.219	1.213
	Interactive	.548	.491	1.116	.264	.985	–
	No highlight	−.983	.478	−2.057*	.040	.859	–
Recall Delayed	Intercept	−1.621	.557	−2.909**	.004	1.701	1.176
	Interactive	.895	.602	1.488	.137	.823	–
	No highlight	−.380	.614	−.619	.536	1.421	–
Recognition Immediate	Intercept	1.200	.379	3.164**	.002	.931	.922
	Interactive	−.129	.391	−.331	.741	.666	–
	No highlight	−1.157	.333	−3.471***	< .001	.664	–
Recognition Delayed	Intercept	.636	.332	1.919	.055	.694	.870
	Interactive	.218	.343	.634	.526	.775	–
	No highlight	−.721	.335	−2.151*	.031	.843	–

Note. Significance level: $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Stimulated Recall Comments

The results from the coding of the stimulated recall transcripts are summarized in Appendix S3 in the Supporting Information. Our analysis revealed that participants noticed the highlighting in both proactive and interactive highlighting conditions. Some participants in the interactive highlighting condition also mentioned that the unanticipated "blinking" was surprising and, at times, distracting. Participants in all three groups reported noticing new words or phrases while reading, with higher noticing rates among those who read unenhanced texts. However, when lexical items were not highlighted, noticing did not necessarily lead to further attempts to reflect on their meanings. In contrast, participants in the highlighting conditions, particularly those exposed to interactive highlighting, more frequently recalled inferring the meanings of unfamiliar lexical items. Additionally, those in the interactive highlighting group more often acknowledged re-encounters with the target collocations.

Different patterns for the groups also emerged regarding their overall reading behaviors. Firstly, participants in the no highlighting group reported pausing to comprehend the text, and those in the highlighting groups tended to recall a previous part of the text. Additionally, those who read unenhanced texts tended to make more comments about fast reading, stating that they perceived the topic as familiar or unimportant, or that fast reading was their habitual way of reading. Another notable trend was that participants who read interactively enhanced texts made substantially fewer comments on re-reading, whereas the other two groups made more references to re-reading. Lastly, there were instances where participants could not recall the specific reasons behind their reading behaviors.

DISCUSSION

This study explored the pedagogical value of gaze-contingent textual enhancement for promoting L2 collocation learning in computer-mediated L2 reading tasks. Three reading conditions were employed, with target collocations either proactively highlighted, interactively highlighted, or not highlighted. Participants' attentional processes were investigated using eye-movement data and stimulated recall comments. The knowledge of the target collocations was assessed using collocation form recall and recognition tests. Now we turn to a discussion of the results addressing each research question.

Impact of Proactive and Interactive Highlighting on Attention to Target Collocations

In this study, attention was explored in terms of the total fixation durations and the total fixation counts recorded on the target collocations. The eye-tracking data revealed that both proactive and interactive highlighting techniques led to more and longer eye fixations on the target collocations. This finding is compatible with previous eye-tracking studies that found positive effects of textual enhancement on visual attention to targeted L2 grammatical constructions (e.g., Indrarathne & Kormos, 2017; Lee & Jung, 2021; Lee & Révész, 2020; Winke, 2013) or lexical items (e.g., Choi, 2017; Godfroid et al., 2013; Jung et al., 2022; Majuddin et al., 2021; Puimège et al., 2021). The eye-movement results are also aligned well with those of Révész et al. (2023), yielding an advantage for gaze-contingent highlighting over no highlighting in promoting visual attention to target linguistic features. As suggested by Révész et al., gaze-contingent highlighting, reacting to learners' eye fixations, seems to have the potential to capture learners' attention during reading effectively. It should be emphasized that participants in this study adopted a task-based approach to the texts, that is, evaluating the suitability of each text for publication in an imagined magazine while filling out a feedback form. That said, this study demonstrates that textual enhancement can maintain its attention-triggering effects when learners process input with a specific goal in their mind, taking a heuristics-oriented and strategic approach to the text (Jung & Yang, 2024). Clearly, the interplay between textual enhancement and task goals warrants more empirical exploration.

The stimulated recall data added insightful information about attention allocation to the target collocations across the three groups. Participants under each condition frequently commented that they had noticed the target collocations. However, the comments of the no highlighting group suggested that participants made no attempts to infer the meaning of the collocations and failed to recognize the target items upon subsequent encounters. In contrast, participants in the highlighting conditions not only reported noticing the highlighted collocations but also mentioned recognizing their repeated encounters during reading. The highlighting groups' comments also implied that they made deliberate attempts to identify the meanings of the target collocations. These comments appear to indicate that the highlighting techniques were successful in boosting participants' attention to the target collocations, prompting them to delve deeper into processing their semantic features. In the interactive highlighting condition, in particular, certain participants reported that they could not help but notice "blinking" during reading. This observation suggests that the dynamic cues delivered in the gaze-contingent mode considerably heightened participants' awareness more than proactive highlighting, leading to more profound engagement with the highlighted collocations.

Impact of Proactive and Interactive Highlighting on Knowledge of the Target Collocation Forms

In this study, the knowledge of the target collocation forms was assessed using collocation recall and recognition posttests. We found that both highlighting techniques helped participants obtain significantly higher scores in the posttests than those under the no highlighting condition. Thus, the findings of this study are consistent with those of previous research (Choi, 2017; Jung et al., 2022; Sonbul & Schmitt, 2013; Szudarski & Carter, 2016; Toomer & Elgort, 2019) that reported positive effects of textual enhancement on L2 collocation learning. Stimulated recall comments, as previously discussed, supported these results, given that participants frequently reported making deliberate attempts to understand the meaning of the highlighted collocations. In other words, highlighting seems to have induced deeper processing of the target collocations and thereby led to the establishment of more durable memory representations. Importantly, only interactive highlighting demonstrated a sustained superior impact on collocation recall to the no highlighting group in the delayed posttest, but with no significant difference observed between interactive and proactive highlighting. It is also notable that a participant in the interactive highlighting group commented finding the unexpected highlighting sometimes distracting, implying a raised awareness. Probably, unanticipated highlighting, sensitively tuned with individual participants' reading pace, effectively captured participants' attention, promoting more attentive processing of the target collocations.

The stronger effects of interactive highlighting may be explained in terms of Wickens's (2015) computational model on visual attention, incorporating the factors salience, effort, expectancy, and value (SEEV). In this model, salience refers to visually noticeable features, and effort involves the muscular effort required to attend to visual stimuli. Also, expectancy relates to the rate of event changes, while value represents the need to process stimuli. When applying this model to the present study, effort and value were held constant for all three groups as learners were exposed to the same texts and performed the same task across all input conditions. Salience seems relevant to the two highlighting conditions in that learners' visual attention was drawn to the target collocations by increasing their typographical salience. In addition, interactive highlighting likely increased the expectancy of events by conditioning highlighting to occur only when learners fixated on the target collocations. That is, learners could not anticipate when highlighting would occur, unlike under the proactive condition, where they could immediately take note of the highlighted collocations when turning to a new page. This unpredictability of gaze-contingent highlighting could have increased the likelihood that the highlighted collocations were noticed. It should be noted, however, that the pronounced effects of gaze-contingent highlighting may also stem from its technological novelty, emphasizing the need for further research into its sustained effectiveness over repeated exposure.

CONCLUSION

The present study aimed to investigate the pedagogical potential of gaze-contingent textual enhancement to promote L2 collocation learning during computer-mediated reading tasks. We found that both highlighting techniques were successful in drawing attention to the target collocations and in improving collocation form recall and recognition posttest scores. Notably, gaze-contingent highlighting had a more lasting impact on the ability to recall collocation forms. Also, the stimulated recall protocols revealed that interactive highlighting was superior at attracting participants' attention to the target collocations. In sum, this study confirmed the facilitative effect of textual enhancement on L2 collocation learning and discovered the value of gaze-contingent textual enhancement as a learner-assistive focus-on-form device for learning collocations.

This study, however, is not free from limitations. First, the sample size and the number of observations (20 participants × 12 items per group) may not offer sufficient statistical power for conducting mixed-effects modeling (Brysbaert & Stevens, 2018), increasing the risk of Type I error. In addition, the target collocations were not tightly controlled in terms of their inherent features, such as Chinese-English congruency or semantic transparency, even though these aspects could potentially affect L2 collocation learning (Sonbul, El-Dakhs, & Alharbi, 2023). Additionally, while participants were carefully selected based on their IELTS or TOEFL scores to guarantee comprehension of the treatment texts, their scores on the Oxford Placement Test were slightly lower than anticipated. Also, despite careful selection and piloting of the target collocations to minimize prior knowledge effects, the possibility of differential prior knowledge cannot be entirely ruled out. Lastly, while the distractor adjectives in the collocation recognition items were meticulously selected, three distractor–noun combinations (*pure carbon*, *urban farm*, and *graphic art*) had MI scores higher than 3.0. While their meanings did not fit in the contexts of the sentences, such high frequencies warrant attention.

Despite the limitations, this study provides several meaningful insights. First, this study showcases the pedagogical potential of the gaze-contingent paradigm (Reder, 1973) in promoting L2 learners' attention to novel collocational features during reading (Révész et al., 2023). While it is true that gaze-contingent highlighting is not currently practical for regular language classrooms, it is important to not only focus on the immediate impacts of readily available technologies but also explore innovative technologies that could positively influence language education in the long term. That said, the comparable effects of gaze-contingent highlighting to that of proactive highlighting, even though it was activated only upon viewing, emphasize the necessity for further exploration in this domain. From a methodological perspective, this study manifests the usefulness of eliciting learners' introspective comments. Specifically, the differential impact of the two highlighting devices emerged more clearly in the stimulated recall comments, compared to the eye fixation indices. Therefore, future studies would also benefit from combining these methods to gain fuller insights into the learners' attentional processes.

ACKNOWLEDGMENT

The work described in this paper was fully supported by a grant from the Research Grants Council of the HKSAR, China (CUHK (24607921)).

THE AUTHORS

Jookyoung Jung is an assistant professor at the Department of English, the Chinese University of Hong Kong. Her research interests include task-based language teaching (TBLT), second language (L2) reading and writing, technology-mediated L2 learning, and the role of individual differences in second language acquisition (SLA). Her recent work has appeared in journals such as IRAL, System, Language Awareness, Language Teaching Research, and Studies in Second Language Acquisition.

Andrea Révész is a professor of Second Language Acquisition at the IOE, University College London. Her research interests lie in second language acquisition, instruction, and assessment, with a particular focus on tasks, input, individual differences, and the neurocognitive processes underlying second language speaking and writing. She serves as editor of the Annual Review of Applied Linguistics and co-editor of the John Benjamins Task-based Language Teaching series. She is past president of the International Association for Task-based Language Teaching.

Matt Stainer is a senior lecturer at Griffith University in Gold Coast, Australia. He completed a BSc Hons in Forensic Psychobiology at Abertay University in Scotland before completing a PhD in Psychology at the University of Dundee (Scotland). His research uses eye tracking and visual attention tasks to examine how people extract visual information from their environments to make decisions.

Ana Pellicer-Sánchez is an associate professor of Applied Linguistics and TESOL at University College London. Her research centres around the teaching and learning of vocabulary in a second/foreign language, with a particular focus on the use of eye-tracking to explore the cognitive processes involved in vocabulary learning. She is co-author of Eye-tracking: A guide for Applied Linguistics Research (CUP) and co-editor of Understanding Formulaic Language (Routledge).

Yoojin Chung is an associate lecturer of Second Language Acquisition at the University of Reading. Her primary research interests lie in instructed second language acquisition, with particular emphases on task-based language teaching and children learning English as a foreign language. She has worked in a wide range of contexts, from teaching English to primary school children; training teachers at international language centres; and working as a TESOL instructor at universities.

Danni Shi is a postdoctoral associate in the Department of Linguistics at Georgetown University, USA. She obtained an MA in TESOL from University College London (UCL) and was awarded a PhD degree in applied linguistics from UCL. Her main research interest falls in the area of second language acquisition, with a particular focus on the roles of tasks and individual differences.

REFERENCES

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmnatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Barton, K. (2015). *MuMIn: Multi-model inference*. R package version 1.13.4. http://cran.r-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R., Singmann, H., & Dai, B. (2014). *Lme4:Linear mixed-effects models using Eigen and S4*. https://cran.r-project.org/package=lme4

Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb–noun collocations. *Language Teaching Research*, *18*, 54–74.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 1–20.

Carrol, G., & Conklin, K. (2014). Eye-tracking multi-word units: Some methodological questions. *Journal of Eye Movement Research*, *7*, 1–11.

Choi, S. (2017). Processing and learning of enhanced English collocations: An eye movement study. *Language Teaching Research*, *21*, 403–426.

Conklin, K., Pellicer-Sanchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge, UK: Cambridge University Press.

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, *75*, 117–139.

Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice* (1st ed.). London, UK: Springer.

Evers, K., Van Belle, G., Steyaert, J., Noens, I., & Wagemans, J. (2017). Gaze-contingent display changes as new window on analytical and holistic face perception in children with autism spectrum disorder. *Child Development*, *89*, 430–445.

Godfroid, A. (2019). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. New York: Routledge.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, *35*, 483–517.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions: An eye-tracking study. *Studies in Second Language Acquisition*, *39*(3), 401–430.

Jung, J., & Lee, M. (2022). Second language reading and recall processes under different reading purposes: An eye-tracking, keystroke-logging, and stimulated recall study. *Language Awareness*, *32*(2), 278–300.
Jung, J., & Révész, A. (2018). The effects of reading activity characteristics on L2 reading processes and noticing of glossed constructions. *Studies in Second Language Acquisition*, *40*(4), 755–780.
Jung, J., Stainer, M., & Tran, M. (2022). The impact of textual enhancement and frequency manipulation on incidental learning of collocations from reading. *Language Teaching Research*. https://doi.org/10.1177/13621688221129994
Jung, J., & Yang, C. L. (2024). The impact of task guidance on incidental collocation learning from task-based reading. *Language Teaching Research*. https://doi.org/10.1177/13621688241270840
Keemink, J. R., Jenner, L., Prunty, J. E., Wood, N., & Kelly, D. J. (2020). Eye movements and behavioural responses to gaze-contingent expressive faces in typically developing infants and infant siblings. *Autism Research*, *14*(5), 973–983.
Lee, H., Kanakogi, Y., & Hiraki, K. (2015). Building a responsive teacher: How temporal contingency of gaze interaction influences word learning with virtual tutors. *Royal Society Open Science*, *2*, 140361.
Lee, M., & Jung, J. (2021). Effects of textual enhancement and task manipulation on L2 learners' attentional processes and grammatical knowledge development: A mixed methods study. *Language Teaching Research*, *28*(4), 1552–1571.
Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition*, *42*, 625–651.
Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York: Routledge.
Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, *65*(1), 185–207.
Little, G. E., Bonnar, L., Kelly, S. W., Lohan, K. S., & Rajendran, G. (2016). Gaze contingent joint attention with an avatar in children with and without. *Joint IEEE International Conference on Development and Learning and Epigenetic Robocins, ICDL-EpiRob 2016*.
Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press.
Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental acquisition of multiword expression through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition*, *43*, 985–1008.
Nation, P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.
Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam, The Netherlands: John Benjamins.
Northbrook, J., & Conklin, K. (2019). Is what you put in what you get out? – Textbook-derived lexical bundles processing in beginner English learners. *Applied Linguistics*, *40*(5), 816–833.
Pellicer-Sánchez, A., Siyanova-Chanturia, A., & Parente, F. (2022). The effect of frequency of exposure on the processing and learning of collocations: A comparison of first and second language readers' eye movements. *Applied PsychoLinguistics*, *43*(3), 727–756.
Peters, E. (2012). Learning German formulaic sequences: The effect of two attention- drawing techniques. *Language Learning Journal*, *40*(1), 65–79.

Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values. *The Modern Language Journal*, *102*(4), 713–731.

Puimège, E., Montero Perez, M., & Peters, E. (2021). Promoting L2 acquisition of multiword units through textually enhanced audiovisual input: An eye-tracking study. *Second Language Research*, *39*(2), 471–492.

Puimège, E., Montero Perez, M., & Peters, E. (2024). The effects of typographic enhancement on L2 collocation processing and learning from reading: An eye-tracking study. *Applied Linguistics*, *45*(1), 88–110.

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Reder, S. M. (1973). On-line monitoring of eye position signals in contingent and noncontingent paradigms. *Behavior Research Methods & Instrumentation*, *5*(2), 218–228.

Révész, A., Stainer, M., Jung, J., Lee, M., & Michel, M. (2023). Using eye-tracking as a tool to develop second language lexical skills. *Language Learning & Technology*, *27*(1), 1–22.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, *35*, 213–235.

Sagarra, N., & Seibert Hanson, A. (2011). Eyetracking methodology: A user's guide for linguistic research. *Hispanic and Lusophone Linguistics*, *4*, 543–555.

Schratz, P. (2021). *oddsratio: Odds ratio calculation for GAM(M)s & GLMs (version 2.0.0)* [R package]. CRAN. https://CRAN.R-project.org/package=oddsratio

Sonbul, S., El-Dakhs, D. A. S., & Alharbi, R. (2023). Rendering natural collocations in a translation task: The effect of direction, congruency, semantic transparency, and proficiency. *International Journal of Applied Linguistics*, *34*(1), 117–133.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge acquisition of collocations under different input conditions. *Language Learning*, *63*, 121–159.

Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, *26*, 245–265.

Text Inspector. (2023). *Analyse the difficulty level of English texts*. https://textinspector.com/

Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning*, *69*, 405–439.

Vernetti, A., Senju, A., Charman, T., Johnson, M. H., Gliga, T., & the BASIS Team. (2020). Simulating interaction: Using gaze-contingent eye-tracking to measure the reward value of social signals in toddlers with and without autism. *Developmental Cognitive Neuroscience*, *20*, 21–29.

Wang, A. F., & Pellicer-Sánchez, A. (2023). Combining eye-tracking and verbal reports in vocabulary research: Benefits and challenges. *Research Methods in Applied Linguistics*, *2*(3), 100063.

Wang, Q., Wall, C. A., Barney, E. C., Bradshaw, J. L., Macari, S. L., Chawarska, K., & Shic, F. (2020). Promoting social attention in 3-year-olds with ASD through gaze-contingent eye tracking. *Autism Research*, *13*, 61–73.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*, 91–120.

Wickens, C. D. (2015). Noticing events in the visual workplace: The SEEV and NSEEV models. In R. Hoffman & R. Parasuraman (Eds.), *Handbook of applied perception* (pp. 749–768). Cambridge, UK: Cambridge University Press.

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogeley, K. (2010). It's in your eyes—Using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *5*(1), 98–107.

Winke, P. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition*, *35*, 323–352.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1: Normality checks of the mixed-effects models.
Appendix S2: True-or-false reading comprehension scores.
Appendix S3: Summary of the coding results.