Szudarski, Pawel and Conklin, Kathy (2014) Short- and long-term effects of rote rehearsal on ESL learners' processing of L2 collocations. TESOL Quarterly, 48 (4). pp. 833-842. ISSN 0039-8322

BRIEF REPORTS AND SUMMARIES

TESOL Quarterly invites readers to submit short reports and updates on their work. These summaries may address any areas of interest to Quarterly readers.

Edited by LIA PLAKANS
University of Iowa

LAWRENCE JUN ZHANG
University of Auckland

Short- and Long-Term Effects of Rote Rehearsal on ESL Learners' Processing of L2 Collocations

PAWEŁ SZUDARSKI
Kazimierz Wielki University
Bydgoszcz, Poland

KATHY CONKLIN
University of Nottingham
Nottingham, England

Worldwide there is thought to be around 750 million people who speak English as a foreign language (Crystal, 2003, p. 69). For these speakers the difference between *make a picture* and *take a picture* may seem arbitrary. However, use of the former is likely to influence how their second language (L2) performance is perceived (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006). Consequently, L2 speakers' use of collocations ("fixed, identifiable, non-idiomatic phrases and constructions"; Benson, Benson, & Ilson, 1997, p. xv) and other formulaic sequences is an important aspect of L2 competence (Wray, 2002).

Several factors appear to influence the acquisition and use of L2 collocations. One of them is L1–L2 collocational congruency. Research has demonstrated that word-for-word translation equivalents (congruent collocations) are processed more efficiently than incongruent collocations (Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010). Moreover, research has demonstrated that collocational frequency, and the frequency of

formulaic sequences more generally, influences processing, with more frequent combinations being processed more quickly (Siyanova-Chanturia, Conklin, & van Heuven, 2011; Wolter & Gyllstad, 2013).

Researchers have also explored the role of different L2 input conditions on the processing of collocations. Sonbul and Schmitt (2013) compared the effects of three treatments (enriched, enhanced, and decontextualized input) on the collocational competence of learners of English as a second language (ESL). For explicit knowledge, they observed an improvement in both receptive and productive tests for all treatment conditions, but for implicit knowledge no gains were found. Peters (2012) examined L2-German learners' acquisition of words and formulaic sequences as dependent on an instructional method (directing learners' attention through instructions) and input enhancement (bolding and underlining). In a form recall test, input enhancement led to gains in learners' knowledge, whereas the instructional method did not seem to affect their results.

The present study brings together a number of questions that have been prominent in the literature by examining the short- and long-term effects of two different input conditions on ESL learners' processing of L2 collocations, as well as exploring the influence of frequency and collocational congruency.

STUDY

The study used two treatments: rote rehearsal (RR) and enhanced rote rehearsal (RR plus). RR was based on previous research (e.g., Hummel, 2010) and relied on the repetition of first language (L1) and L2 collocations. RR plus was the same treatment but included input enhancement, in the form of underlining, which aimed at increasing the graphical salience of collocations. The study tested the following hypotheses:

1. RR and RR plus improve ESL learners' processing of L2 collocations.
2. RR plus leads to greater gains than RR.
3. Irrespective of method, frequent collocations are processed more efficiently than infrequent collocations.

METHOD

Participants

Twenty-six (8 males and 18 females; age: $M = 27$; $SD = 5.17$) L1-Polish participants living in the United Kingdom were recruited for

TABLE 1
TABLE 1
Learners' Mean Vocabulary Levels Test Scores Across Frequency Levels

	2,000	3,000	Academic	5,000	10,000
Mean	29.96	28.12	28.96	26.08	14.58
SD	1.75	2.14	1.11	4.42	7.33

Note. The maximum score for each level is 30.

the study. They were all advanced learners of English, as demonstrated by the results of Schmitt, Schmitt, and Clapham's (2001) Vocabulary Levels Test (see Table 1).

Instrumentation

The study included three sessions: a pretest, a treatment with an immediate posttest, and a delayed posttest. Eighteen participants were tested individually via an acceptability judgment task (Yamashita & Jiang, 2010) before and after the treatment. Eight learners formed a control group, as did an additional 14 native English speakers. These two groups only took part in the judgment task.

Materials

The experiment included 80 collocations: 40 adjective-noun (AN; e.g., *warm welcome*) and 40 verb-noun (VN; e.g., *hold a raffle*) combinations. They were divided into two counterbalanced lists in such a way that each of them contained 20 frequent (10 AN and 10 VN) and 20 infrequent (10 AN and 10 VN) combinations (see online supporting information).

All collocations were carefully matched, such that they all had similar length and high mutual information scores indicating strongly connected combinations (for a discussion of mutual information, see Manning & Schütze, 1999). With regard to collocational frequency, frequent combinations occurred 79 times or more in the British National Corpus (BNC) ($M = 380$), while infrequent collocations occurred 30 times or less ($M = 11$). Additionally, the frequency of the individual words of the collocations was controlled. All nouns in the frequent combinations belonged to the first 3,000 most frequent words in the BNC ($M = 6,612$), whereas all nouns in the infrequent collocations represented low frequency levels ($M = 198$). This was done to ensure that the two sets were sufficiently different from each other.

Procedure

DMDX (Forster & Forster, 2003) was used to measure participants' reaction times (RTs) and error rates (ERs). First, a fixation point was presented on a computer screen for 500 ms. Then, a collocation was displayed for 5000 ms and participants had to decide whether it was an acceptable phrase in English by pressing either a *yes* or *no* button. They were told to do this as quickly and accurately as possible.

Ten practice items were included before the experiment started. Participants were presented with 120 items: 40 target collocations and 80 fillers (e.g., *quick glass*, *give development*). The presence of these non-collocates ensured that the test was meaningful and that responses were not biased to a *yes* or *no*.

Sonbul and Schmitt (2013) establish the need for careful sequencing of measures so that earlier tests do not contaminate later results. Thus, the treatment was conducted 2 weeks after the pretest to minimize practice effects. In the treatment itself, the target collocations were embedded in 40 short (10–12 words) sentences that contained only active, present-tense forms. Each sentence was displayed on the screen in two language versions and learners were asked to read both versions aloud. They had three seconds for each Polish sentence and six seconds for the English ones. Importantly, in the RR plus condition, all collocations were underlined as in the example below. This type of enhancement has been successfully used in Peters's (2012) research on formulaic language.

Example: Training stimuli, with collocation underlined in the RR plus condition.

Ta firma nie <u>odnosi zysku</u> z usług, które oferuje.
This company does not <u>make a profit</u> from the services it offers.

The whole treatment lasted about 12 min and ended with a short language background questionnaire. Next, participants were given the immediate posttest. Following Schmitt's (2010) suggestion that a delay of 3 weeks should be indicative of durable learning, the delayed posttest was administered after 6 weeks.

RESULTS

Reaction Times (RTs)

Mean RTs are reported in Table 2. All incorrect responses were excluded from the RT analysis and analysed separately. Outliers longer

Mean Reaction Times (in Milliseconds) on Four Types of Collocations Tested During Three Sessions

	Session	Frequent AN M (SD)	Infrequent AN M (SD)	Frequent VN M (SD)	Infrequent VN M (SD)
RR plus	Pretest	1,210 (334)*	1,851 (644)*	1,185 (308)	1,776 (899)
N = 18	Posttest	963 (209)	1,450 (528)	1,096 (319)	1,435 (600)
	Delayed posttest	1,115 (413)	1,311 (464)*	1,202 (325)	1,331 (337)*
RR	Pretest	1,170 (267)*	1,722 (556)*	1,214 (338)	1,695 (731)*
N = 18	Posttest	990 (188)	1,492 (537)	1,135 (336)	1,367 (445)
	Delayed posttest	1,112 (383)	1,526 (562)	1,146 (265)	1,363 (358)
Control	Pretest	1,380 (390)	1,976 (488)	1,414 (363)*	1,659 (412)
N =8	Posttest	1,173 (378)	1,668 (721)	1,200 (360)	1,505 (589)
	Delayed posttest	1,135 (385)*	1,446 (534)*	1,235 (192)	1,380 (375)*
NSs		925 (167)	1,202 (245)	913 (119)	1,070 (196)
N = 14					

Note. $*p < .05$.

than 5000 ms were excluded. Using this cutoff point, 4.71% of the pretest, 1.34% of the posttest, and 0.76% of the delayed posttest data were excluded. Because the data were not normally distributed, nonparametric tests were used. First, Wilcoxon tests comparing learners' RTs on the pretest, posttest, and delayed posttest were conducted. Both treatments resulted in significant differences between the pretest and posttest results for the frequent AN (RR plus: $z = -3.28$; $p = .001$; $r = .78$; RR: $z = -2.98$; $p = .003$; $r = .70$) and infrequent AN collocations (RR plus: $z = -2.11$; $p = .04$; $r = .51$; RR: $z = -2.72$; $p = .006$; $r = .64$). As far as the VN collocations are concerned, only the RR condition reached significance on the infrequent items ($z = -2.44$; $p = .02$; $r = .59$). For comparisons between the pretest and the delayed posttest data, only the RR plus condition revealed significant differences on the infrequent AN ($z = -2.77$; $p = .006$; $r = .68$) and infrequent VN collocations ($z = -2.07$; $p = .04$; $r = .49$). However, these results did not differ from the control group where RTs also decreased, which means that no claims about long-term effects of the treatment can be made.

Second, given that both RR plus and RR led to shorter RTs on the AN collocations, their effectiveness was compared. No significant differences between the two conditions were found: (frequent AN: $z = -1.07$; $p = .27$; $r = .25$; infrequent AN: $z = -0.92$; $p = .36$; $r = .22$). Moreover, Mann-Whitney tests were run to compare learners and NSs. The pretest results showed that NSs were significantly faster than learners in both the RR plus condition (frequent AN: $U = 62$; $z = -2.43$; $p = .02$; $r = .43$; infrequent AN: $U = 39$; $z = -3.18$; $p = .001$; $r = .56$; frequent VN: $U = 58$; $z = -2.58$; $p = .01$; $r = .46$; infrequent VN: $U = 35$;

$z = -3.46$; $p = .001$; $r = .61$) and the RR condition (frequent AN: $U = 58$; $z = -2.58$; $p = .01$; $r = .46$; infrequent AN: $U = 54$; $z = -2.74$; $p = .006$; $r = .48$; frequent VN: $U = 60$; $z = -2.51$; $p = .01$; $r = .44$; infrequent VN: $U = 37$; $z = -3.26$; $p = .001$; $r = .58$). Immediately after the treatment, these differences disappeared in both the RR plus and RR conditions. The only exception was the infrequent VN collocations in the RR condition where NSs were still significantly faster than learners ($U = 72$; $z = -2.05$; $p = .04$; $r = .36$). The delayed posttest revealed that the knowledge of the VN collocations was not retained—learners were again significantly slower than NSs in both the RR plus condition (frequent VN: $U = 54$; $z = -2.74$; $p = .006$; $r = .48$; infrequent VN: $U = 66$; $z = -2.28$; $p = .02$; $r = .40$) and the RR condition (frequent VN: $U = 45$; $z = -3.08$; $p = .002$; $r = .53$; infrequent VN: $U = 67$; $z = -2.24$; $p = .03$; $r = .40$).

Third, Wilcoxon tests were conducted to compare participants' RTs on the frequent and infrequent collocations. During all the sessions, the former were processed significantly more quickly than the latter in both the RR plus (pretest VN: $z = -3.57$; $p = .001$; $r = .08$; pretest AN: $z = -3.37$; $p < .001$; $r = .87$; delayed VN: $z = -2.03$; $p = .04$; $r = .48$) and RR (pretest VN: $z = -2.86$; $p = .004$; $r = .07$; pretest AN: $z = -3.59$; $p < .001$; $r = .85$; delayed VN: $z = -3.38$; $p = .001$; $r = .80$; delayed AN: $z = -3.29$; $p = .001$; $r = .80$) conditions. The only exception was the delayed posttest data on the AN collocations in the RR plus condition, which failed to reach significance ($z = -1.85$; $p = .06$; $r = .44$).

Error Analysis

As Table 3 demonstrates, error levels on the pretest were relatively low and comparable for learners and NSs. Thus, there was little room for improvement resulting from the treatment. Despite apparent gains on the infrequent collocations, no claims about learners' improved accuracy can be made, because the experimental groups did not differ significantly from the control group that also improved: RR plus (AN: $U = 66$; $z = -0.43$; $p = .67$; $r = .08$; VN: $U = 49.5$; $z = -1.53$; $p = .13$; $r = .03$) and RR (AN: $U = 55.5$; $z = -1.33$; $p = .18$; $r = .26$; VN: $U = 65.5$; $z = -0.4$; $p = .69$; $r = .08$).

When the experimental conditions were compared, no significant differences were found. The only exception to this was in the immediate posttest data, where the infrequent VN collocations elicited significantly fewer errors in the RR condition compared to the RR plus condition ($z = -2.12$; $p = .03$; $r = .51$). However, this improvement was short term, as the delayed posttest data did not reveal this difference ($z = -0.91$; $p = .37$; $r = .21$).

TABLE 3
Mean Error Rates on Four Types of Collocations Tested During Three Sessions

Group	Session	Frequent AN M (SD)	Infrequent AN M (SD)	Frequent VN M (SD)	Infrequent VN M (SD)
RR plus	Pretest	9.8 (0.6)	6.9 (3.3)*	9.3 (1.4)	6.4 (2.4)*
N = 18	Immediate posttest	9.9 (0.5)	9.1 (1.7)	9.9 (0.5)	9.1 (1.0)
	Delayed posttest	9.9 (0.5)	9.2 (1.7)*	9.7 (1.0)	9.4 (1.8)*
RR	Pretest	9.8 (0.6)	6.5 (2.5)*	9.6 (0.9)	7.1 (2.4)*
N = 18	Immediate posttest	10.0 (0)	9.3 (1.0)	9.7 (1.0)	9.5 (1.3)
	Delayed posttest	9.6 (0.9)	9.7 (1.0)*	9.6 (0.9)	9.1 (1.2)*
Control	Pretest	9.9 (0.4)	7.1 (2.1)*	9.6 (0.7)	7.1 (2.1)*
N = 8	Immediate posttest	10.0 (0)	8.6 (2.0)	9.8 (0.5)	8.6 (1.8)
	Delayed posttest	9.9 (0.3)	9.5 (0.8)*	9.6 (0.5)	8.9 (1.6)*
NS		9.9 (0.3)	9.4 (3.3)	9.9 (0.4)	8.5 (2.0)
N = 14					

Note. *$p < .05$. One point awarded for each correct answer (max score = 10).

Frequency of collocations was found to influence the error rate. The pretest data revealed that in both conditions learners' errors for the infrequent collocations were significantly higher than for the frequent collocations: RR plus (AN: $z = -2.88$; $p = .004$; $r = .70$; VN: $z = -3.22$; $p = .001$; $r = .76$) and RR (AN: $z = -3.25$; $p = .001$; $r = .77$; VN: $z = -3.07$; $p = .002$; $r = .72$). When the same comparisons were made after the treatment, the delayed posttest data revealed that these differences were no longer found: RR plus (AN: $z = -1.6$; $p = .11$; $r = .38$; VN: $z = -0.82$; $p = .41$; $r = .19$) and RR (AN: $z = -0.45$; $p = .66$; $r = .11$; VN: $z = -1.63$; $p = .10$; $r = .38$).

DISCUSSION AND CONCLUSIONS

This study examined ESL learners' processing of AN and VN collocations following two different instructional treatments: RR and RR plus. Both treatments led to more effective processing of collocations, as demonstrated by faster RTs for the frequent and infrequent AN collocations. The fact that VN collocations did not demonstrate an overall benefit from the treatments indicates differences in L2 learners' processing of different types of collocational patterns. This is likely to have been caused by L1–L2 congruency effects (30 out of 40 VN vs. 1 out of 40 AN were incongruent collocations). Previous research (Wolter & Gyllstad, 2011; Yamahita & Jiang, 2010) has demonstrated the difficulty of processing incongruent items by L2 learners. Our study seems to suggest that short repeated presentation, even with highlighting, is not enough to enhance L2 learners' performance on such items.

Moreover, it should be stressed that the positive effects of the instruction seem to be short term, as the delayed posttest data revealed a decline in the participants' knowledge (higher RTs in the delayed posttest results). Thus, it appears that L2 collocational knowledge should be consolidated and recycled soon after it has been acquired if it is to be retained for longer periods of time. We second Durrant and Schmitt's (2010) call for research on the role of recycling in L2 lexical development. It is important to establish when it should be introduced in teaching programs so that learners are able to enjoy its long-term benefits. Sonbul and Schmitt's (2013) study, with post-tests 2 weeks after their treatment, found that learners retained collocations at the explicit levels of mastery. Our study showed that with a short treatment (repetition of 40 sentences), L2 collocational knowledge had declined after 6 weeks. This suggests that the consolidation of such knowledge should take place 3–5 weeks after instruction; otherwise attrition will be evident.

Furthermore, similar gains in both the RR and RR plus groups seem to indicate that the repetition of L1 and L2 sentences containing the target collocations was the most important factor in facilitating learners' processing of them. It is likely that the use of both languages (L1-Polish and L2-English) contributed to the benefits of the treatment. Importantly, the lack of additional gains in the RR plus group shows that when there are repetitions and presentation of the L1 and L2, there is no further benefit of underlining. This is an interesting finding that betters our understanding of input enhancement, which might be effective only in combination with other types of treatment (Szudarski, 2012).

Finally, the study demonstrates the importance of collocational frequency. We found significant differences between learners' processing of the frequent and infrequent collocations, with the former resulting in shorter RTs and higher accuracy rates. This adds evidence to a growing body of research (e.g., Siyanova-Chanturia et al., 2011) showing that L2 learners' processing of formulaic language is sensitive to frequency effects. Moreover, the accuracy analysis revealed that the L2 learners had achieved native-like levels for the frequent collocations. This shows that such combinations do not pose difficulty for ESL learners at a receptive level of mastery, which accords with Laufer and Waldman's (2011) assertion that "the real problem of collocations in L2 does not lie in recognition but in learning to use them properly" (p. 652).

In sum, the study provides empirical evidence for the short-term benefits of rote rehearsal in L2 learners' processing of collocations. Repetition was found to facilitate ESL learners' results, with the addition of input enhancement (underlining) not resulting in more gains.

The study also revealed that learners' gains were not retained after 6 weeks, which emphasizes the importance of consolidating the newly acquired L2 collocational knowledge. Lastly, the frequency and type of collocations were confirmed as two important factors that influence L2 learners' lexical development.

THE AUTHORS

Paweł Szudarski is a lecturer at Kazimierz Wielki University, Bydgoszcz, Poland. Recently he completed his PhD studies at the University of Nottingham. Earlier he worked for the Center for Applied Linguistics in Washington, DC. His research interests include multiword units and phraseology, L2 vocabulary acquisition and testing, English language learners, and corpus linguistics.

Kathy Conklin is an associate professor in psycholinguistics at the University of Nottingham. She is interested in multiword units, automatic word activation in an L1 vs. L2, bilingual/monolingual cognitive control, and implicit L2 learning. To investigate L1/L2 language representation and processing, she makes use of behavioral measures, eye-tracking, and EEG.

REFERENCES

Benson, M., Benson, E., & Ilson, R. (Eds.). (1997). *The BBI combinatory dictionary of English: A guide to word combinations.* Amsterdam, the Netherlands: John Benjamins.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10,* 245–261. doi:10.1191/1362168806lr195oa

Crystal, D. (2003). *English as a global language.* Cambridge, England: Cambridge University Press.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26,* 163–188. doi:10.1177/0267658309349431

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35,* 116–124. doi:10.3758/BF03195503

Hummel, K. M. (2010). Translation and short-term L2 vocabulary retention: Hindrance or help? *Language Teaching Research, 14,* 61–74. doi:10.1177/1362168809346497

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61,* 647–672. doi:10.1111/j.1467-9922.2010.00621.x

Manning, C. D., & Schűtze, H. (1999). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press.

Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *Language Learning Journal, 40,* 65–79. doi:10.1080/09571736.2012.658224

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Basingstoke, England: Palgrave Macmillan.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing, 18,* 55–89. doi:10.1177/026553220101800103

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37,* 776–784. doi:10.1037/a0022531

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning, 63,* 121–159. doi:10.1111/j.1467-9922.2012.00730.x

Szudarski, P. (2012). *The effectiveness of input enhancement and input enrichment in the acquisition of formulaic sequences by EFL learners.* Paper presented at the 22nd Annual Conference of the European Second Language Acquisition, Adam Mickiewicz University, Poznań, 5-8 September.

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics, 32,* 430–449. doi:10.1093/applin/amr011

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition, 35,* 451–482. doi:10.1017/S0272263113000107

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge, England: Cambridge University Press.

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly, 44,* 647–668. doi:10.5054/tq.2010.235998

Supporting Information

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

Table S1: Target Items.