ISSUES IN LANGUAGE TEACHING
2026, VOL. 15, NO. 1, 57–79
https://doi.org/10.22054/ilt.2026.87016.927

ALLAMEH TABATABA'I UNIVERSITY PRESS

OPEN ACCESS

Input Scheduling and L2 Grammar Learning: Investigating the Delayed Effect of Spacing

Reza Zabihi [1]✉, Mahnaz Khosrojerdi [2], and Mehdi Mehranirad [3]

1. *Corresponding Author,* Department of English, University of Neyshabur, Neyshabur, Iran. E-mail: zabihi@neyshabur.ac.ir
2. Department of English, University of Neyshabur, Neyshabur, Iran. E-mail: kmahnaz235@gmail.com
3. Department of English, University of Neyshabur, Neyshabur, Iran. E-mail: mehrani@neyshabur.ac.ir

ABSTRACT

The timing and distribution of instructional input play a critical role in shaping language learning outcomes. This study examined the effects of input spacing on second language (L2) grammar learning by comparing massed and distributed presentation formats. Twenty Persian-speaking pre-intermediate learners at a private language institute in Mashhad, Iran, participated in the study. The participants were taught four unfamiliar grammatical rules across two instructional sessions. Each learner encountered two rules under a massed input condition (two sessions separated by a two-day interval) and two rules under a distributed input condition (two sessions separated by a nine-day interval), allowing for a within-subjects comparison. A delayed posttest was administered four weeks after the final instructional session to assess long-term retention. The results of the t-test revealed a significant advantage (M = 6.95, SD = 1.98) for the distributed input condition (9-day interval) over the massed condition (M = 5.70, SD = 2.17), $t(19) = 2.190$, $p < .05$, suggesting that increasing the lag between instructional encounters can enhance long-term retention of L2 grammatical knowledge. These findings support the spacing effect in L2 acquisition and offer pedagogical insights into how input scheduling can influence the retention of grammatical knowledge in foreign language instruction. The study contributes to the growing body of research advocating for the strategic timing of instructional input to maximize long-term learning outcomes in L2 education.

ARTICLE TYPE:
Research Article

ARTICLE HISTORY:
Received 19 November 2025
Received in revised form 20 May 2026
Accepted 12 June 2026
Published Online 30 June 2026

KEYWORDS:
distribution of practice, input scheduling, L2 grammar learning, retention, spacing effect

CITE THIS ARTICLE: Zabihi, R., Khosrojerdi, M., & Mehranirad, M. (2026). Input Scheduling and L2 Grammar Learning: Investigating the Delayed Effect of Spacing. *Issues in Language Teaching*, *15*(1), 57-79. https://doi.org/10.22054/ilt.2026.87016.927

Introduction

One important factor that can influence second language (L2) learning is the time interval over which learning occurs. This interval can take the form of either massed practice (i.e., when learning is concentrated into a single session or across sessions with minimal temporal lag) or distributed practice (i.e., when learning is spread across multiple sessions separated by longer intersession intervals) (Rogers, 2017; Rohrer, 2015). A growing body of research has examined the differential effects of massed versus distributed study sessions (e.g., Amini et al., 2021; Baleghizadeh & Asadi, 2013; Cepeda et al., 2006; Dempster, 1988, 1996; Rohrer & Pashler, 2007). Across a wide range of domains, including mathematics (Rohrer & Taylor, 2006), vocabulary learning (Bahrick et al., 1993; Bahrick & Phelps, 1987; Bloom & Shuell, 1981; Dempster, 1987; Glenberg & Lehmann, 1980), reading comprehension (Reder & Anderson, 1982), biology concepts (Reynolds & Glaser, 1964), and name–face association (Landauer & Bjork, 1978), findings have consistently demonstrated the superiority of distributed over massed practice. This phenomenon is commonly referred to as the *spacing effect* (Rogers, 2017; Rohrer, 2015). Thalheimer (2006) defined the spacing effect as the enhancement of learning that occurs when material is presented more than once, with a time interval between repetitions. Regardless of how the repetitions are presented, learning tends to be more effective when there is spacing between exposures.

Another factor that plays a critical role in L2 learning is task repetition. This is particularly important because developing automaticity requires learners to use language knowledge effortlessly, which, in turn, demands repeated exposure and practice. Theoretical models (McLaughlin, 1987; Skehan, 1998) and pedagogical approaches grounded in cognitive psychology (Hadley & Reiken, 1993; Skehan, 1998) highlight repetition as a key mechanism for enabling learners to move from controlled language use toward more spontaneous and automatic language production. However, an important practical question remains: how should L2 practice schedules be organized to promote proceduralization and automatization? In other words, when is the optimal time to repeat instructional input? As early as the 19th century, Ebbinghaus (1885) found that reviewing material at spaced intervals significantly improved retention.

Following the recognition that spaced practice is generally more effective than massed practice, researchers began to explore the optimal *intersession interval* (ISI)—the time between study sessions (Rogers & Cheung, 2018). Determining the ideal ISI has proven to be a complex challenge. Crowder

(1976) proposed that the optimal ISI depends on the *retention interval* (RI)—the time between the end of practice and the test. Glenberg (1976) similarly suggested that the relationship between ISI and RI follows an inverted U-shaped curve: as RI increases, the optimal ISI also increases, but if the ISI becomes too long, retention declines. Cepeda et al. (2006), in their meta-analysis of 184 studies, supported this conclusion and reaffirmed the U-shaped relationship. These findings suggest that there is a specific ISI that maximizes long-term retention for a given RI, and that excessively long ISIs may lead to forgetting.

Most prior research on the spacing effect has been conducted in highly controlled laboratory settings (Bird, 2010; Cepeda et al., 2006; Rogers, 2015; Rohrer, 2015; Serrano, 2011). Therefore, there is a pressing need to conduct more studies in ecologically valid learning environments. *Ecological validity* refers to the extent to which research findings can be generalized to real-world contexts where the studied phenomena naturally occur (Schmuckler, 2001). The present study addresses this gap by examining the impact of input spacing on some target L2 grammatical structures using training procedures that reflect authentic classroom conditions.

The acquisition of grammatical knowledge is a central component of L2 proficiency, as it underlies learners' ability to produce structurally accurate and meaningful utterances in communication. Moreover, given that grammar learning often requires repeated encounters with target forms to facilitate proceduralization, it provides a suitable domain for investigating how the temporal distribution of instructional input influences long-term retention.

Literature Review

Theoretical Foundations and Early Models of the Spacing Effect

The spacing effect is a fundamental concept in experimental psychology that illustrates how learning is more effective when study sessions are spaced apart rather than massed together. This robust finding has attracted interest not only from researchers but also from educators and learners seeking to optimize instructional time and study habits. Since its first systematic documentation by Ebbinghaus (1885), the spacing effect has been replicated across diverse populations, materials, and learning contexts, and is now considered one of the most reliable phenomena in human memory (Cepeda et al., 2006; Kang, 2016; Weinstein et al., 2018).

Theoretical frameworks explaining the spacing effect have evolved over time, starting with the all-or-none theory proposed by Estes et al. (1960).

According to this theory, a given item on a trial is either completely learned or not at all. This binary view led to the prediction that longer gaps between study sessions would always enhance retention because longer gaps increase the likelihood that initial learning has been forgotten, necessitating re-learning and thus strengthening memory traces. Consequently, the theory implied that there should be no "optimal" interval—longer gaps would simply be better.

However, subsequent empirical findings have challenged the all-or-none theory's predictions. In reality, excessively long intervals can lead to forgetting to the point where re-learning becomes inefficient, suggesting an optimal balance must exist. This led to the development of alternative theoretical accounts, including deficient-processing theory, encoding variability theory, and study-phase retrieval theory, which emphasize cognitive effort, contextual variability, and retrieval processes as mechanisms underlying spacing benefits (Delaney et al., 2010; Kang, 2016; Maddox, 2016). More recently, the reconsolidation framework has also been proposed, suggesting that spaced retrieval strengthens memory by repeatedly reactivating and stabilizing memory traces (Antony et al., 2017; Pan & Rickard, 2018).

One practical application of spacing principles emerged in 1972 with the introduction of the Leitner System, a spaced repetition algorithm designed to optimize flashcard study by increasing intervals between reviews based on learner performance. This system was revolutionary for self-directed learning, especially before the advent of computers and allowed learners to focus more on items they struggled with while spacing out easier items. Although popular and effective to some extent, learners often reported dissatisfaction with the accuracy of Leitner-based systems in reflecting their true knowledge states (Settles & Meeder, 2016). More recently, adaptive spaced-repetition systems using computational models and artificial intelligence, such as SuperMemo and Anki, have improved prediction of memory decay and individualized scheduling, demonstrating enhanced learning efficiency compared to fixed schedules (Lindsey et al., 2014; Tabibian et al., 2019).

Cognitive psychology research has since focused on mathematically modeling the relationship between the ISI and the RI, or the time between initial study and final test. These models have converged on the finding that the optimal ISI is approximately 10–30% of the RI. In other words, the spacing between study sessions should be roughly a fraction of the total time over which retention is desired. Recent meta-analyses have confirmed this proportional relationship while also emphasizing that optimal spacing varies

depending on learner characteristics, task complexity, and instructional context (Carpenter et al., 2022; Küpper-Tetzel, 2021). This insight has been influential in educational practice, guiding the timing of reviews and practice sessions for maximal retention. However, an important question remains: Does this ratio hold true across domains, especially for complex tasks like second language acquisition?

Unlike well-defined tasks often used in experimental psychology, such as memorizing word lists, L2 grammar acquisition—the dependent variable in the present study—involves the development of both explicit knowledge of grammatical rules and the ability to use those rules accurately in communication. Bird (2010) highlighted this distinction by noting that psychological studies typically employ narrowly defined memory tasks, whereas language learning requires learners to integrate grammatical form, meaning, and use. This distinction is directly relevant to spacing research because the effectiveness of different spacing intervals—the independent variable in this study—may depend on the type and complexity of the knowledge being learned. For example, learning grammatical structures such as *used to* or *wish + would*, which were examined in this study, requires not only remembering the rule but also retrieving and applying it accurately during language production. This process involves both declarative and procedural knowledge systems. Recent applied linguistics research has emphasized that these two types of knowledge may respond differently to distributed practice schedules, with proceduralized grammar knowledge particularly sensitive to the timing of repeated input and retrieval opportunities (Li & DeKeyser, 2021; Suzuki & Sunada, 2021). Therefore, investigating the effect of spacing intervals on L2 grammar acquisition is essential for determining whether general spacing principles derived from simple memory tasks apply equally to complex grammatical learning in classroom contexts.

Empirical Studies on Spacing in L2 Grammar Learning

Several empirical studies have sought to apply and test spacing effects in the context of L2 grammar learning, producing a nuanced and sometimes contradictory picture. Overall, findings suggest that spacing benefits L2 grammar learning, but optimal intervals vary depending on instructional conditions, linguistic targets, and learner variables (Rogers & Cheung, 2018; Suzuki et al., 2019).

Bird (2010) conducted a study with 38 Malay university English learners focusing on distinguishing English verb tenses: simple past, present perfect,

and past perfect. Participants were assigned to two groups with different spacing schedules (three-day vs. 14-day ISI) and tested at short-term (7 days) and long-term (60 days) RIs. The study used worksheets containing sets of sentences, where learners identified and corrected verb tense errors. The results showed that while short-term retention did not differ significantly between groups, the 14-day ISI group demonstrated better long-term retention, supporting the idea that longer spacing benefits long-term memory.

Rogers (2015) expanded this line of research with a study on complex syntactic structures involving 100 grammatical and ungrammatical sentences. Learners were divided into groups practicing either with two-day or seven-day intervals. Posttests included an immediate test and a delayed test after 42 days. Similar to Bird's findings, no difference appeared in immediate posttests, but the seven-day ISI group outperformed the two-day group in delayed retention, reinforcing the benefit of longer spacing for durable learning. These findings align with classical spacing theories but raise questions about short-term versus long-term learning mechanisms.

In contrast, Suzuki and DeKeyser (2017) investigated whether spacing effects apply similarly to proceduralized grammar knowledge, such as the production of Japanese verb forms. Their study with beginner learners compared one-day and seven-day ISIs using a series of oral production tasks and posttests measuring accuracy and response speed. Surprisingly, both groups performed similarly on delayed posttests in terms of accuracy, but the one-day ISI group was faster in a picture description task, indicating possible advantages of shorter spacing for automatization of production skills. This diverged from Bird's and Rogers' results and suggested that spacing effects might differ depending on the type of linguistic knowledge (declarative vs. procedural).

Suzuki (2017) conducted a replication and extension study comparing 3-day and 7-day ISIs with morphological rules. Sixty participants engaged in four training sessions designed to promote oral production under either a shorter spaced (3-day interval) or longer spaced (7-day interval) learning condition. From the third training session onward, learners in the 3-day interval group began to produce target rules more accurately than those in the seven-day interval group, and this advantage was maintained on both 7- and 28-day delayed posttests, although no significant differences were observed in production speed. These findings contrast with prior spacing research and highlight the potential variability in optimal intersession intervals depending on task demands and learner-related factors.

Extending this line of classroom-based research, Suzuki and Hanzawa (2022) investigated the effects of task repetition under massed, short-spaced,

and long-spaced practice schedules in an English as a foreign language (EFL) instructional context. Learners performed the same oral narrative task six times either consecutively within a single class session (massed practice), distributed within a single class (short-spaced), or across two classes separated by a one-week interval (long-spaced). Immediate posttest results based on a novel narrative task indicated that massed practice led to greater reductions in breakdown fluency; however, it was also associated with slower articulation rates and increased verbatim repetition, suggesting potential trade-offs in fluency development. Furthermore, the effects of repetition schedules were limited on a one-week delayed posttest involving a novel task, although massed practice led to greater reliance on verbatim repetition when the practiced task was repeated. These findings highlight the complex and potentially task-dependent effects of repetition schedules in classroom-based L2 learning contexts.

Kasprowicz et al. (2019) further examined spacing effects in an ecologically valid classroom context with young English learners of French. Groups practiced inflectional verb morphology at seven-day and 3 day intervals using a game-based app, alongside a control group with regular instruction. Posttests measured sentence-picture matching, grammaticality judgment, and language analytic ability (LAA). Results showed minimal differences between spacing conditions, but individual learner differences, especially LAA, moderated learning success. This suggests that in authentic classroom settings, learner characteristics may have greater influence than spacing schedules alone.

More recent research has continued to refine these findings. For example, Suzuki and Sunada (2021) found that spaced practice significantly improved grammar production accuracy compared to massed practice, particularly at delayed posttests. Similarly, Li and DeKeyser (2021) demonstrated that distributed practice facilitated long-term proceduralization of grammatical knowledge, although shorter spacing sometimes benefited initial skill acquisition. Furthermore, Rogers and Cheung (2018) reported that optimal spacing intervals differed depending on whether learning outcomes were measured through receptive or productive tasks, highlighting the importance of aligning spacing schedules with instructional goals.

In addition, meta-analytic evidence suggests that spacing effects in L2 learning are robust but moderate in size and are influenced by instructional design, feedback, and learner proficiency levels (Kang et al., 2024; Suzuki et al., 2019). These findings indicate that spacing interacts dynamically with pedagogical and learner variables rather than functioning as a universally fixed principle.

Methodological Considerations and Limitations

Despite these valuable insights, the reviewed studies face several methodological challenges that limit the generalizability of their findings. Bird's (2010) study, for example, used explicit error detection worksheets focusing narrowly on verb form accuracy, which does not fully capture spontaneous grammatical production in real communication. Full language proficiency entails the ability to produce grammatically accurate sentences at discourse levels with little conscious attention, something these tasks do not measure. Similarly, Rogers's (2015) grammaticality judgment tests assess recognition rather than productive competence, leaving open questions about transfer to actual language use.

Laboratory-based settings, like those in Suzuki and DeKeyser's (2017) studies, while highly controlled, may not reflect classroom dynamics where learners receive instruction from teachers and interact socially. Kasprowicz et al.'s (2019) classroom study incorporated real-world settings but minimized teacher involvement, relying heavily on digital tools, thus limiting ecological validity in contexts where teacher guidance is crucial. Recent reviews have emphasized the need for more classroom-based longitudinal studies to better understand how spacing functions in authentic educational environments (Küpper-Tetzel, 2021; Rogers & Cheung, 2018).

Another limitation concerns the RIs used in these studies. Many employed intervals either very short (e.g., 7 days) or very long (e.g., 42 days or 9 weeks), which may not align with typical institutional testing schedules, such as the 30-day or one-semester exams common in many educational systems including Iran's. This mismatch raises questions about the practical applicability of these findings in standard educational settings.

Finally, most studies did not sufficiently account for individual differences among learners, such as cognitive abilities, motivation, prior knowledge, or language analytic skills, even though evidence from Kasprowicz et al. (2019) suggests these factors significantly influence learning outcomes. Ignoring individual variability can obscure the nuanced effects of spacing and reduce the precision of instructional recommendations. Recent research has increasingly emphasized learner-adaptive approaches, suggesting that optimal spacing may vary across individuals depending on proficiency level, working memory capacity, and learning rate (Li & DeKeyser, 2021; Lindsey et al., 2014).

Taken together, while spacing effects are generally beneficial for L2 grammar learning, their optimal application depends on multiple factors including task type, linguistic skill targeted, learning context, timing of

assessments, and individual learner differences. Future research should seek to incorporate these variables in more ecologically valid designs to better inform classroom practices. In particular, further investigation is needed to determine optimal spacing schedules aligned with real classroom timelines and assessment cycles, especially in under-researched educational contexts.

Purpose of the Study

Although previous studies have examined the spacing effect in grammar learning, questions remain about the generalizability of their findings to typical classroom settings across different language institutes. Notably, the practice schedules used in those studies often do not align with the actual schedules followed in many institutions. Furthermore, the instructional methods employed were not representative of the conventional grammar teaching approaches used by classroom teachers. To enhance the ecological validity of this research, the present study adopts an RI consistent with the schedules of the language institutes from which data were collected, and grammar is taught using the conventional classroom methods practiced by teachers.

Another limitation in prior research is that learners typically participated in either spaced-short or spaced-long learning conditions exclusively. Given that individual differences may influence outcomes, this study addresses this issue by having participants engage in both spaced-short and spaced-long learning conditions. This design aims to reduce the confounding effects of individual learner differences on the results. Although Rogers and Cheung (2018) attempted to overcome this limitation in their study on vocabulary learning, their findings cannot be directly generalized to grammar acquisition, as grammar learning is inherently more complex than memorizing word lists (Wulf & Shea, 2002). Additionally, there is evidence that the benefits of distributed learning vary depending on the complexity of the task (Donovan & Radosevich, 1999).

Given these considerations, the present study investigates the spacing effect specifically in the context of second language grammar learning. It seeks to answer the following research questions:

1) Is there a significant difference between the effects of shorter versus longer intersession intervals on learners' acquisition of the target grammatical structures?

2) To what extent do shorter versus longer intersession intervals influence learners' learning and retention of the target L2 grammatical structures?

Method

Participants

The study involved 20 non-native English speakers (13 females, seven males) enrolled at a private language institute in Mashhad, located in northeastern Iran. Their ages ranged from 21 to 30 years, with a mean age of 22.87 years. All participants were native Persian speakers and had no proficiency in any language other than Persian (L1) and English as their L2. Based on the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001), they were classified as pre-intermediate (A2) learners. This classification was determined through a combination of a placement test and a semi-structured oral interview conducted by the language institute.

Specifically, the placement test included 40 multiple-choice questions assessing grammar and vocabulary knowledge (e.g., verb tense selection, modal verbs, and sentence completion), and learners scoring between 21 and 30 were classified as A2. The oral interview lasted approximately five–seven minutes and required learners to describe past experiences, daily routines, and future plans, which confirmed their ability to produce simple sentences but not more complex grammatical structures. This two-stage assessment procedure ensured accurate proficiency classification and is consistent with standard L2 research practices (Li & DeKeyser, 2021; Suzuki & DeKeyser, 2017).

All participants consented to study four grammatical rules over two learning episodes, which were administered under two different spaced learning conditions: spaced-short (two-day interval) and spaced-long (9-day interval). Participants were informed that the purpose of the sessions was instructional, and the experimental manipulation of spacing was not disclosed to prevent expectancy effects.

All 20 participants who completed the pretest also participated in the immediate and delayed posttests. Therefore, no attrition occurred between testing sessions, and the final sample size for all statistical analyses remained constant (N = 20).

Instruments

Four grammatical rules unfamiliar to the learners were carefully selected through a collaborative process between the researcher and the participating teachers. This ensured that the target grammar points were indeed new to the learners. To verify unfamiliarity, learners were informally asked prior to

instruction to provide example sentences using each target structure. When learners were unable to produce correct examples (e.g., incorrect use of *wish + would* or confusion between *must* and *should*), the structure was confirmed as suitable for inclusion.

In the first learning session, two rules were taught: *used to/didn't use to* and so, *neither + auxiliaries* (restricted to the simple present), or alternatively *used to/didn't use to* and *might*. In the second session, the other two rules were introduced: the difference between *must* and *should*, and the structure after *wish* (specifically *wish + would/wouldn't*) or *so, neither + auxiliaries* (again limited to the simple present), along with expressions of movement. Instruction followed a structured three-phase procedure:

1) Presentation phase (10 minutes): The teacher explained the rule using examples (e.g., "*I used to play football.*" vs. "*I didn't use to like coffee.*").
2) Controlled practice phase (10 minutes): Learners completed sentence-completion exercises (e.g., "*When I was younger, I _____ walk to school every day.*").
3) Production phase (10 minutes): Learners produced original sentences orally and in writing (e.g., describing childhood habits using used to).

These rules were chosen as they represent grammatical structures just beyond the pre-intermediate proficiency level of the learners. Learning was assessed using a delayed posttest administered four weeks after instruction. The test format mirrored the standard end-of-semester assessments used by the language institute, ensuring ecological validity for participants.

The posttest consisted of 20 items divided into two sections: Section 1: Sentence completion (10 items); Example: "*When I was a child, I _____ (use to / used to) play outside every day.*" Section 2: Sentence production (10 items); Example prompt: "Write one sentence using '*wish + would*' about something in your life."

Each correct response received one point, resulting in a maximum score of 20. Production items were scored using a predefined rubric focusing on grammatical accuracy. To ensure content validity, the posttest items were reviewed by two experienced EFL instructors with more than five years of teaching experience at the institute. They confirmed that the test items adequately reflected the target grammatical structures taught during the instructional sessions and were appropriate for learners at the A2 proficiency level. In addition, the production items were independently scored by two trained raters using the predefined rubric. Inter-rater reliability was calculated using Cohen's kappa coefficient, which indicated substantial agreement between raters ($\kappa = .89$). Any discrepancies were resolved through discussion.

Finally, participants were asked whether they had engaged in additional practice outside of class related to the target grammatical structures. Data from those who reported extra practice were excluded from further analysis to maintain the integrity of the findings. This control procedure is consistent with recommendations in spacing research, as external exposure may confound distributed practice effects (Cepeda et al., 2006; Kang, 2016).

Procedure

This study employed a within-subjects design to investigate whether, and to what extent, input spacing affects the learning of the target L2 grammatical structures. The within-subjects approach was chosen to control for individual differences in learning styles and abilities. The independent variable—input provision interval—was manipulated by having the same group of L2 English learners participate in two different treatment conditions: spaced-short and spaced-long intervals. The dependent variable was the learners' grammar performance, measured by posttests administered after each treatment condition.

The teaching and review schedule was based on the principle that the optimal ISI should fall within 10-30% of the RI. The study lasted six weeks. During the first week, learners attended a 30-minute session to learn half of the target grammatical items. In the second week, another 30-minute session introduced the remaining items, followed by a 20-minute review session two days later (ten minutes devoted to each half). Finally, four weeks after the completion of the course, a 20-minute posttest was administered to assess retention. The following table summarizes the schedule of teaching, review, and testing sessions.

Table 1.
Experimental design

	Monday	Tuesday	Wednesday
Week 1	Training session 1 (items 1-2)		
Week 2	Training session 2 (items 3-4)		Review session (items 1-4)
Week 3			
Week 4			
Week 5			
Week 6			28-day RI test

Before the experiment began, the researcher conducted a pre-briefing session with the teachers involved in the study. During this meeting, the purpose of the study, the specific grammar rules to be taught, and the experimental

procedures were thoroughly discussed and agreed upon. It was decided that in the first teaching session, two grammar rules—used to/didn't use to and so, neither + auxiliaries (for the simple present), or alternatively, used to/didn't use to and might—would be taught over approximately thirty minutes. In the second teaching session, the remaining two grammar rules—the difference between *must* and *should*, and the structure after *wish* (specifically *wish + would/wouldn't*), or *so, neither + auxiliaries* (for the simple present), and expressing movement—would be covered in another thirty-minute session.
The review session, lasting twenty minutes, followed the same teaching methods but involved less time for explanations and fewer practice activities for each grammar rule to reinforce learning efficiently. The test administered four weeks after the initial session was designed to resemble the language institute's regular final tests to maintain ecological validity and reduce any unfamiliarity with the assessment format.

Results

The primary research question in this study was whether there is a significant difference between the effects of spaced-short and spaced-long learning schedules on the acquisition of some target L2 grammatical structures. Table 2 presents the descriptive statistics for learners' grammar scores under both spaced-short and spaced-long learning conditions.

Table 2.
Descriptive statistics: grammar scores across conditions

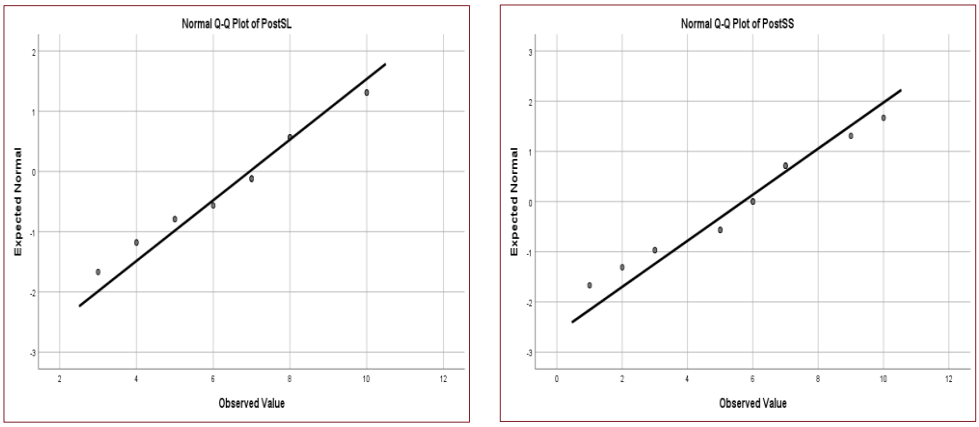
		Grammar scores	
	N	***M***	***SD***
Learning schedule			
Spaced-long	20	5.70	2.17
Spaced-short	20	6.95	1.98

In addition, inspection of individual scores revealed that 14 out of 20 learners performed better under the spaced-short condition, suggesting a consistent pattern favoring shorter spacing. Next, the learners' post-treatment grammar scores for the two learning schedules were compared. Given the small sample size (N = 20), it was necessary to check the normality of the data distribution before conducting a t-test, as this assumption is especially important for samples smaller than 25. To assess normality, the Shapiro-Wilk (SW) test was employed. The Shapiro-Wilk test is similar in purpose to the

Kolmogorov-Smirnov (KS) test, but it is often preferred due to its higher statistical power. Results of the Shapiro-Wilk test indicated no significant deviation from normality for either condition (spaced-long, p = .180; spaced-short, p = .154), confirming that the data met the assumptions required for the t-test (see Figure 1).

Figure 1.

Normal Q-Q plots of grammar scores across the two learning schedules



Using an alpha level of .05, a paired samples (repeated measures) t-test was conducted to examine whether there was a significant difference between the mean post-treatment grammar scores in the two learning conditions. The results revealed a statistically significant difference between the spaced-long instruction (Condition 1; M = 5.70, SD = 2.17) and the spaced-short instruction (Condition 2; M = 6.95, SD = 1.98), $t(14) = -2.190$, $p < .05$ (two-tailed). The mean difference in grammar scores between the two conditions was 1.25, with a 95% confidence interval ranging from -2.44 to -0.055 (see Table 3).

Table 3.

Paired samples t-test results for grammar scores across learning conditions

Pair	Mean Difference	SD	SE	95% CI Lower	95% CI Upper	t	df	p
Spaced-Short – Spaced-Long	-1.25	2.55	0.57	-2.44	-0.055	-2.19	19	.041

While the paired t-test indicated a significant difference between the mean grammar scores of the two learning conditions, it does not provide information about the magnitude of this difference. To further evaluate the

robustness of the findings, effect size analysis showed Cohen's d = 0.49, indicating a moderate effect. This magnitude is comparable to effect sizes reported in prior L2 spacing research (Li & DeKeyser, 2021; Suzuki & Sunada, 2021). These findings demonstrate that the spaced-short condition resulted in approximately 22% higher performance compared to the spaced-long condition, indicating meaningful pedagogical impact.

Discussion

This study set out to explore how the temporal spacing between learning episodes influences the acquisition of the target L2 grammatical structures in a naturalistic classroom context. The findings indicated that the spaced-short learning schedule (with a two-day intersession interval) led to significantly better grammar retention after four weeks compared to the spaced-long schedule (9-day interval). The moderate effect size further underscores the practical importance of these results, suggesting that optimizing the timing of input provision can meaningfully enhance grammar learning outcomes. Importantly, this finding demonstrates that not all distributed schedules are equally effective and that excessively long intervals may weaken retrieval success and reduce consolidation efficiency, particularly for learners at lower proficiency levels (Carpenter et al., 2022; Kang, 2016).

These results corroborate the spacing effect documented extensively in cognitive psychology, which suggests that distributed practice enhances long-term retention better than massed or widely spaced practice (Cepeda et al., 2006; Donovan & Radosevich, 1999). Specifically, the finding that the two-day interval approximates 10% of the RI aligns with the model proposed by Cepeda and colleagues, which identifies an optimal ISI as roughly 10–30% of the total RI. This provides empirical support for applying cognitive principles of memory consolidation to L2 grammar instruction. From a theoretical perspective, these findings are particularly consistent with study-phase retrieval theory, which proposes that spacing enhances learning when the interval is long enough to make retrieval effortful but short enough to allow successful retrieval (Carpenter et al., 2022; Kang, 2016). In the present study, the two-day interval likely created desirable retrieval difficulty, strengthening memory representations, whereas the 9-day interval may have exceeded learners' retrieval threshold, resulting in weaker reinforcement.

Furthermore, this study's findings resonate with prior research in second language acquisition by Suzuki and DeKeyser (2017) and Suzuki (2017), who reported superior performance with shorter spacing in sentence production tasks. The similarities in instructional method — initial explanation followed

by production-focused practice — and the use of production-based assessments may account for the converging results. This suggests that the spacing effect may be particularly robust when learners actively produce language rather than passively recognize it, supporting the hypothesis that the complexity of grammar learning requires appropriately timed, repeated engagement to consolidate procedural knowledge. This interpretation is also supported by recent classroom-based research showing that shorter distributed intervals promote proceduralization of grammar by increasing opportunities for retrieval and restructuring of linguistic representations (Li & DeKeyser, 2021; Suzuki & Sunada, 2021). Similarly, Suzuki and Hanzawa (2022) found in an EFL classroom context that although massed task repetition improved certain aspects of immediate fluency, its benefits were less robust on delayed measures involving novel tasks, suggesting that distributed practice may be more conducive to durable learning outcomes.

In contrast, previous studies such as Bird (2010), Rogers (2015), and Kasprowicz et al. (2019) have yielded mixed or contradictory findings regarding input spacing effects in grammar learning. The present study's within-subjects design, where each learner experienced both spacing conditions, helped control for individual differences — a factor often overlooked in between-subjects designs. This methodological distinction likely contributed to clearer detection of spacing effects by reducing inter-individual variability. Moreover, the discrepancy between the present findings and those reporting advantages for longer spacing (e.g., Bird, 2010) may be explained by differences in RIs, task demands, and learner characteristics. For example, longer spacing intervals tend to benefit learning when RIs are very long or when learners have higher proficiency levels (Carpenter et al., 2022; Cepeda et al., 2008), whereas shorter spacing may be more effective for relatively short retention goals and developing learners, as in the present study.

Moreover, differences in the nature of instruction and assessment across studies may explain divergent findings. For instance, Bird (2010) and Rogers (2015) used recognition-based grammar tasks, whereas this study employed production tasks that arguably demand deeper processing and stronger memory encoding. Kasprowicz et al. (2019) implemented app-based grammar teaching, which contrasts with the authentic classroom instruction here involving direct teacher explanations and interactive practice. These contextual differences highlight the necessity of considering ecological validity and task type when interpreting spacing effects, as cognitive processes engaged during learning and retrieval vary considerably depending on these factors (Kasprowicz et al., 2019; Suzuki & DeKeyser, 2017). This

finding supports claims that spacing effects may be stronger for productive, effortful language use than for passive recognition, as production requires more elaborate retrieval and integration processes (Carpenter et al., 2022; Suzuki & Sunada, 2021).

The study also draws attention to the importance of individualized learning conditions in L2 classrooms. By having all participants engage in both spaced-short and spaced-long conditions, this research minimized confounding effects of learner variability such as motivation, cognitive styles, and prior knowledge, which often complicate interpretations in between-groups designs. This strengthens the argument that spacing effects are robust and generalizable within typical classroom populations. More importantly, this study contributes to the field by providing ecologically valid evidence from an authentic instructional setting, addressing longstanding calls for classroom-based spacing research rather than laboratory-based experiments (Küpper-Tetzel, 2021).

Nonetheless, it is important to acknowledge limitations. The study was limited to a relatively small sample of pre-intermediate adult learners, restricting generalization to other proficiency levels or age groups. The focus on grammar also leaves open the question of whether similar spacing benefits occur for other language domains, such as vocabulary acquisition or pragmatic skills. Additionally, the study design did not explore underlying cognitive mechanisms, such as attentional allocation or retrieval practice effects, that may mediate the spacing effect in L2 learning. Future studies incorporating multiple posttests and cognitive process measures could clarify how spacing interacts with retrieval effort and proceduralization processes (Li & DeKeyser, 2021).

From a pedagogical perspective, the findings advocate for scheduling grammar review sessions relatively soon after initial teaching — approximately two days later — to maximize learning efficiency. This may challenge conventional practices that space grammar lessons more widely or cluster practice into massed sessions. Teachers and curriculum designers should consider integrating shorter, well-timed review opportunities to support durable grammar acquisition, particularly in intensive language programs where optimizing learning within limited contact hours is critical. More specifically, the findings suggest that teachers can improve learning outcomes without increasing instructional time simply by reorganizing review timing. This has important implications for curriculum design, syllabus planning, and digital learning platforms, where review schedules can be systematically optimized based on cognitive principles (Carpenter et al., 2022; Lindsey et al., 2014).

Beyond pedagogy, this study makes three key theoretical and empirical contributions. First, it provides direct classroom-based evidence supporting study-phase retrieval theory in L2 grammar acquisition. Second, it helps resolve conflicting findings in prior L2 spacing research by demonstrating that shorter spacing may be optimal under certain instructional and retention conditions. Third, it extends spacing effect research beyond laboratory settings by demonstrating its applicability in authentic language classrooms, thereby strengthening the ecological validity of cognitive learning theories.

In sum, this research provides valuable evidence supporting the spacing effect in authentic L2 grammar learning contexts and underscores the need for carefully designed instructional schedules grounded in cognitive science. Future research should extend these findings by examining larger and more diverse learner populations, incorporating multiple language skills, and probing the cognitive processes underlying spacing effects. Such investigations will enhance our theoretical understanding and inform best practices for language education worldwide.

Conclusion and Implications

This study contributes to the growing body of research demonstrating the impact of input spacing on L2 grammar acquisition, highlighting that shorter-spacing intervals yield better learning outcomes in classroom settings. By implementing a within-subjects design and maintaining ecological validity through authentic teaching methods and schedules, the study offers valuable evidence that grammar practice sessions spaced two days apart improve retention compared to longer intervals.

Limitations include the relatively small sample size (N = 20) and the restriction to pre-intermediate adult learners, limiting generalizability. Additionally, the study did not explore the effects of input spacing across different proficiency levels or age groups. Future research should address these limitations by involving larger and more diverse participant pools, examining various proficiency stages, and considering child learners. Further studies could also broaden the scope to include other language skills beyond grammar, such as vocabulary and the four main language skills (speaking, listening, reading, and writing). Moreover, research exploring the optimal timing for repeated grammar practice and investigating why shorter spacing enhances learning would deepen theoretical understanding and inform teaching practice. In summary, while confirming the benefits of spaced-short learning intervals, this study encourages continued exploration of spacing

effects in diverse educational contexts to better tailor language instruction for maximum effectiveness.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the author used ChatGPT 4 in order to improve readability and language of the work. After using this tool/service, the author reviewed and edited the content as needed and take full responsibility for the content of the published article.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. However, copyright and ethical restrictions may apply to the availability of the data.

Disclosure Statement

The authors declare that they have no competing financial or non-financial interests in relation to the work described in this manuscript.

Ethics Approval

The research was ethically approved by the Ethics Committee of the University of Neyshabur. The study was conducted in accordance with the guidelines governing research involving human participants, as outlined by the Ethics Committee of the University of Neyshabur.

Prior to their participation, participants were informed of the purpose and the procedure of the study, and informed consent was obtained from all individual participants included in the study.

Funding

This study received no financial support from any agencies or academic sectors.

ORCID

Reza Zabihi https://orcid.org/0000-0001-7210-5424
Mahnaz Khosrojerdi https://orcid.org/0009-0005-1397-8512
Mehdi Mehranirad https://orcid.org/0000-0001-6825-2042

References

Amini, B., Bayat, A. & Mahmoodi, K. (2021). The Impact of Sequencing Repeated Familiar Tasks on Listening Performance. *Issues in Language Teaching*, *10*(2), 325-350. https://doi.org/10.22054/ilt.2022.61019.597

Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. Trends in Cognitive Sciences, 21(8), 573–576. https://doi.org/10.1016/j.tics.2017.05.001

Bahrick, H. P., & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 344–349. https://doi.org/10.1037/0278-7393.13.2.344

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, *4*(5), 316–321. https://doi.org/10.1111/j.1467-9280.1993.tb00571.x

Baleghizadeh, S., & Asadi, R. (2013). The effect of task repetition and task recycling on EFL learners' oral performance. *Issues in Language Teaching*, *2*(2), 137–163.

Bird, S. (2010). Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics*, *32*(3), 435–452. https://doi.org/10.1017/S0142716409990343

Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research, 74*(4), 245–248. https://doi.org/10.1080/00220671.1981.10885372

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2022). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review*, *34*(2), 987–1016. https://doi.org/10.1007/s10648-021-09621-8

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Crowder, R. G. (1976). *Principles of learning and memory*. Lawrence Erlbaum Associates.

Delaney, P. F., Verkoeijen, P. P., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation*, *53*, 63–147. https://doi.org/10.1016/S0079-7421(10)53003-2

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology*, *79*(2), 162–170. https://doi.org/10.1037/0022-0663.79.2.162

Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, *43*(8), 627–634. https://doi.org/10.1037/0003-066X.43.8.627

Dempster, F. N. (1996). Distributing and managing the conditions of encoding practice. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 318–339). Academic Press.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795–805. https://doi.org/10.1037/0021-9010.84.5.795

Ebbinghaus H. (1885). *Memory: A contribution to experimental psychology*. Teachers College Press.

Estes, W. K., Hopkins, B. L., & Crothers, E. J. (1960). All-or-none and conservation effects in the learning and retention of paired associates. *Journal of Experimental Psychology, 60*(4), 329–339. https://doi.org/10.1037/h0049974

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior, 15*(1), 1–16. https://doi.org/10.1016/S0022-5371(76)90006-5

Glenberg, A. M., & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory & Cognition, 8*(6), 528–538. https://doi.org/10.3758/BF03198283

Hadley, A. O., & Reiken, E. (1993). *Teaching language in context, and teaching language in context--workbook*. Heinle & Heinle Publishers.

Kang, S. H. K. (2016). Spaced repetition promotes efficient and effective learning: Policy implications for instruction. *Educational Psychology Review*, 28(1), 3–16. https://doi.org/10.1007/s10648-015-9335-4

Kasprowicz, R., Marsden, E., & Sephton, N. (2019). Investigating distribution of practice effects for the learning of foreign language verb morphology in the young learner classroom. *The Modern Language Journal, 103*(3), 580–606. https://doi.org/10.1111/modl.12563

Küpper-Tetzel, C. E. (2021). Understanding the distributed practice effect: Strong effects and open questions. *Educational Psychology Review*, 33(2), 753–768. https://doi.org/10.1007/s10648-020-09552-4

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). Academic Press.

Li, S., & DeKeyser, R. M. (2021). Implicit language aptitude: Conceptualizing the construct, validating the measures, and examining the evidence: Introduction to the special issue. *Studies in Second Language Acquisition*, *43*(3), 473–497. https://doi.org/10.1017/S0272263121000024

Lindsey, R. V., Shroyer, J. D., Pashler, H., & Mozer, M. C. (2014). Improving students' long-term knowledge retention through personalized review. *Psychological Science*, *25*(3), 639-647. https://doi.org/10.1177/0956797613504302

Maddox, G. B. (2016). Spacing, repetition, and memory. In J. H. Byrne (Ed.), Learning and memory: A comprehensive reference (2nd ed., Vol. 2, pp. 277–294). Academic Press.

McLaughlin, B. (1987). *Theories of second language acquisition*. Arnold.

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*(7), 710–756. https://doi.org/10.1037/bul0000151

Reder, L. M., & Anderson, J. R. (1982). Effects of spacing and embellishment on memory for the main points of a text. *Memory & Cognition, 10*(2), 97–102. https://doi.org/10.3758/BF03197675

Reynolds, J. H., & Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology*, *55*(5), 297–308. https://doi.org/10.1037/h0044227

Rogers, J. (2015). Learning second language syntax under massed and distributed conditions. *TESOL Quarterly*, *49*(4), 857–866. https://doi.org/10.1002/tesq.229

Rogers, J. (2017). The spacing effect and its relevance to second language acquisition. *Applied Linguistics*, *38*(6), 906–911. https://doi.org/10.1093/applin/amw042

Rogers, J., & Cheung, A. (2018). Input spacing and the learning of L2 vocabulary in a classroom context. *Language Teaching Research*, *24*(5), 616–641. https://doi.org/10.1177/1362168817717667

Rohrer, D. (2015). Student instruction should be distributed over long time periods. *Educational Psychology Review*, *27*(4), 635–643. https://doi.org/10.1007/s10648-015-9334-x

Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Current Directions in Psychological Science*, *16*(4), 183–186. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology*, *20*(9), 1209–1224. https://doi.org/10.1002/acp.1266

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02

Serrano, R. (2011). The time factor in EFL classroom practice. *Language Learning*, *62*(1), 117–145. https://doi.org/10.1111/j.1467-9922.2011.00659.x

Settles, B., & Meeder, B. (2016). A trainable spaced repetition model for language learning. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 1848–1858). https://doi.org/10.18653/v1/P16-1171

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning*, *67*(3), 512–545. https://doi.org/10.1111/lang.12226

Suzuki, Y., & DeKeyser, R. M. (2017). Effects of distributed practice on the proceduralization of morphology. *Language Teaching Research*, *21*(5), 662–688. https://doi.org/10.1177/1362168816653101

Suzuki, Y., & Hanzawa, K. (2022). Massed task repetition is a double-edged sword for fluency development: An EFL classroom study. *Studies in Second Language Acquisition*, *44*(2), 536–561. https://doi.org/10.1017/S0272263121000358

Suzuki, Y., & Sunada, M. (2021). Optimizing the spacing of practice for second language grammar learning: Evidence from classroom research. *Language Teaching Research*, *25*(6), 981–1003. https://doi.org/10.1177/1362168820903345

Suzuki, Y., Nakata, T., & DeKeyser, R. (2019). Optimizing second language practice in the classroom: Perspectives from cognitive psychology. *Language Learning*, *69*(Suppl. 1), 187–223. https://doi.org/10.1111/lang.12335

Tabibian, B., Upadhyay, U., De, A., Zarezade, A., Schölkopf, B., & Gomez-Rodriguez, M. (2019). Enhancing human learning via spaced repetition optimization. *Proceedings of the National Academy of Sciences*, *116*(10), 3988–3993. https://doi.org/10.1073/pnas.1815156116

Thalheimer, W. (2006). Spacing learning events over time: What the research says. http://www.work-learning.com/download/Spacing.pdf

Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. Cognitive Research: *Principles and Implications*, *3*(2). https://doi.org/10.1186/s41235-017-0087-y

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review, 9*(2), 185–211. https://doi.org/10.3758/BF03196276

APPENDIX A. Post-test

Time: 25–30 minutes

Section 1: *Complete each sentence using the correct form. (1 point each)*

1. When I was a child, I _____ (use to / used to) play outside every day.
2. She didn't _____ (use to / used to) like vegetables, but now she loves them.
3. I am tired. _____ am I. (So / Neither)
4. He doesn't like coffee. – _____ do I. (So / Neither)
5. You _____ study more if you want to pass the exam. (must / should)
6. You _____ wear a helmet when riding a motorcycle. It's the law. (must / should)
7. I _____ go to the party tonight, but I'm not sure. (might / must)
8. She _____ be at home now; the lights are off. (might / should)
9. I wish my neighbor _____ stop making noise at night. (would / will)
10. I wish it _____ rain so much in this city. (would / will)

Section 2: Write ONE correct sentence for each prompt. (1 point each)

11. Write a sentence about a past habit using **used to**.
12. Write a sentence about something you did NOT do in the past using **didn't use to**.
13. Write a sentence using **so + auxiliary** (agreement in the present).
14. Write a sentence using **neither + auxiliary** (agreement in the present).
15. Write a sentence giving advice using **should**.
16. Write a sentence expressing obligation using **must**.
17. Write a sentence using **might** to show possibility.
18. Write a sentence using **wish + would** about something in your life.
19. Write a sentence describing movement (e.g., go, come, walk, run).
20. Write a sentence using **might** about possibilities of future transportation.