Memory & Cognition
https://doi.org/10.3758/s13421-023-01461-z

Sleep-dependent consolidation effects on foreign language word acquisition in a virtual reality environment

Cong Liu[1] · Yanna Mao[1] · Xiaohan Wang[1] · John W. Schwieter[2,3] · Lu Jiao[1]

Accepted: 23 August 2023

Abstract

Sleep-dependent consolidation is important for novel word learning, but previous studies have neglected the potential modulating role of learning environments. The present study examines sleep-dependent consolidation effects by comparing learning in a virtual reality (VR) environment and in a traditional picture-word (PW) environment. Two groups of Chinese–English bilinguals were randomly assigned to a VR or PW environment. In both learning environments, they learned novel words in Korean, a language with which they had no prior experience. All participants learned one set of novel words on Day 1 and another set on Day 2. An explicit recognition task and an implicit primed lexical-decision task were employed to measure sleep-dependent consolidation effects from the two environments. Results revealed sleep-dependent consolidation effects in both explicit and implicit measures, but only the primed lexical-decision task showed an influence of learning environment, suggesting that novel words learned via VR had better consolidation. Taken together, our findings suggest that a VR environment that fosters a rich sensory experience facilitates sleep-dependent consolidation effects. We argue that these results provide new evidence and implications for the complementary learning system (CLS) model.

Keywords Novel word acquisition · Sleep-dependent consolidation · Semantic priming · Virtual reality

Word acquisition is crucial for foreign-language learning (Hacking & Tschirner, 2017; Tartaro et al., 2021), given that it is fundamentally linked to other linguistic performance, including grammatical and general communicative abilities (Barcroft, 2004; Kern, 1989; Proctor et al., 2005; Smidt & Hegelheimer, 2004). However, word acquisition is a challenge for many learners, a likely reason for why researchers and educators are interested in how to best address it (Lan et al., 2015; Li & Jeong, 2020). A key factor facilitating novel words into lexical networks is sleep-dependent consolidation, which helps to integrate novel words with existing ones (Liu & van Hell, 2020; Tartaro et al., 2021). The idea that sleep is important to memory consolidation and neural plasticity is long standing (Graves, 1936; Klinzing et al., 2016). Despite accumulated evidence examining the role of sleep-dependent consolidation on word acquisition, studies have been limited to traditional learning environments through picture–word/word–word association paired learning. It remains unclear as to how sleep-dependent consolidation affects novel word acquisition in a (simulated) realistic learning environment, an issue tested in the present study.

Background

The theoretical motivation of the present study is the complementary learning system (CLS) model (Davis & Gaskell, 2009; McClelland et al., 1995). The CLS model emphasizes that there are two stages of word acquisition—namely, the initial familiarization stage and the slow lexical consolidation stage (Davis & Gaskell, 2009; Kumaran et al., 2016). In the rapid initial familiarity stage, novel words are hypothesized to be quickly encoded into episodic memory through the hippocampus system and related regions of the medial temporal lobe, allowing learners to recall novel words

✉ Lu Jiao
jiaolu902@126.com

[1] Department of Psychology, Normal College & School of Teacher Education, Qingdao University, NingXia R, No.308, Qingdao 266071, China

[2] Language Acquisition, Multilingualism, and Cognition Laboratory/Bilingualism Matters @ Wilfrid Laurier University, Waterloo, Canada

[3] Department of Linguistics and Languages, McMaster University, Hamilton, Canada

Published online: 04 October 2023

immediately. In the slow lexical consolidation stage, off-line consolidation transforms hippocampus memory representations to stable and long-term neocortical representations and allows novel words to interact with existing ones. Overall, according to the CLS model, off-line consolidation may be a central feature of novel word acquisition which contributes to the integration of novel words into existing lexical networks (Liu & van Hell, 2020; Palma & Titone, 2021).

Empirical evidence has supported the indispensable role of off-line consolidation on novel word acquisition (Bakker et al., 2015; Dumay & Gaskell, 2007; Lindsay & Gaskell, 2013; Szmalec et al., 2012). To investigate off-line consolidation, the primed lexical-decision task, which combines a semantic decision with a primed task, is often used to measure novel word acquisition (Liu & van Hell, 2020). It is typically found that responses to target words (e.g., “cat”) that are preceded by semantically related prime words (e.g., “dog”) yield shorter response times and higher accuracy compared with words with semantically unrelated prime words (e.g., “pen”), an observation typically referred to as the semantic priming effect (e.g., Lei et al., 2022; Liu & van Hell, 2020; Meyer & Schvaneveldt, 1971). For example, in a study by Tamminen and Gaskell (2013), two groups of participants were required to learn novel words and were subsequently tested by an explicit definition recall task and an implicit primed lexical-decision task. The difference between the two groups was the extent of off-line consolidation: For one group, there was a time interval for off-line consolidation between learning and testing, and for the other group, there was no time interval. The results of the primed lexical-decision task revealed that participants responded faster to semantically related trials than semantically unrelated trials only for these words learned with off-line consolidation time, suggesting the importance of off-line consolidation in word acquisition.

Despite some studies revealing that off-line consolidation effects can be triggered without sleeping between learning and testing (Lindsay & Gaskell, 2013; Palma & Titone, 2021), a widely accepted viewpoint is that sleep is important for off-line consolidation effects (Bakker et al., 2015; Kurdziel et al., 2017; Kurdziel & Spencer, 2016; McClelland et al., 1995; Tham et al., 2015; Wang et al., 2017). For instance, in a study by Dumay and Gaskell (2007), two groups of participants learned novel words and were then tested at three time points. A “morning group” learned novel words in the morning and was required to complete the first test immediately after learning, a second test 12 hours after learning (i.e., on the same day without sleep), and a third test 24 hours after learning (i.e., the next day after having slept). The second group, “the evening group,” received the same learning in the evening immediately followed by a test, and a second and third test (12 and 24 hours later, respectively) were completed the following day (i.e., after having slept). Behavioral results showed that despite the absence of a consolidation effect in the immediate test for both groups, the evening group showed off-line consolidation effects in both delay tests, whereas the morning group failed to observe the consolidation effect after the 12-hour period without sleep. Similar findings were reported by Kurdziel and Spencer (2016), whose results supported the function of sleep-dependent consolidation effects on novel word acquisition. Taken together, these findings suggest that compared with wakefulness, sleep is an important and effective way to examine off-line consolidation effects. When individuals are in a state of sleep, there is no need for their brain to receive and process new stimulus, thus providing the opportunity to allocate more cognitive resources to consolidate learning and establish memory (McClelland et al., 1995).

However, further neuroscientific evidence for sleep-dependent consolidation has indicated that novel words after consolidation cannot function as stable lexical representations compared with existing words (Bakker et al., 2015; Lei et al., 2022; Liu & van Hell, 2020). For instance, Bakker and colleagues (2015) required participants to learn a set of novel words on Day 1 (Set 1, remote condition) and another set on Day 2 (Set 2, recent condition). Immediately after the learning on Day 2, participants were administered a definition recall-task, a comprehension-based recognition task, and a primed lexical-decision task to measure learning outcomes and consolidation effects. Behavioral and electrophysiological data revealed that for the remote condition, sleep-dependent consolidation after learning contributed to the semantic processing of novel words, but lexical retrieval was more strategic and less automatic than for existing words. Later, Liu and van Hell (2020) replicated the same design of Bakker et al. (2015), but added a retest on Day 8 to increase the off-line consolidation period. The findings also revealed that novel words learned after sleep-dependent consolidation were not fully integrated into the semantic network compared with existing words. Thus, sleep-dependent consolidation may be a gradual and protracted process (see also Tartaro et al., 2021).

One potential explanation for slower integration processes is that sleep-dependent consolidation studies have been limited to traditional learning environments, such as word– word pairs, word–picture pairs, and word–definition association learning (Lei et al., 2022; Liu & van Hell, 2020). The levels-of-processing theory emphasizes that enriched learning environments (e.g., novel words learned in the real world or learned with bodily involvement) boost learning because encoding and semantic integration involve several levels (Craik & Lockhart, 1972; Li & Jeong, 2020). According to the levels-of-processing theory, when learners in traditional learning environments memorize isolated novel words by focusing on their spelling and translation, the processing mechanisms involved are shallow and consequently make semantic integration occur at a slow space. In a (simulated) realistic environment, on the contrary, which involves a rich sensory experience, learners can establish a direct link between novel words and their concepts by interacting with target objects within a contextualized environment. This learning environment may lead to deeper processing that facilitates semantic integration and consolidation effects (Lan et al., 2015; Li & Jeong, 2020). The potential effects of learning environment on word acquisition and consolidation could be theoretically supported by embodied cognition theory which emphasizes that whole-body interactions positively shape experience and knowledge (Barsalou, 2008). However, for many foreign language learners, acquiring novel words in a realistic environment, such as being in the country in which the target language is spoken, is unattainable.

Recently, the development of virtual reality (VR) technology, with immersive involvement and interaction characteristics, provides an opportunity to simulate realistic learning environments. In a VR environment, learners can hear novel words, interact with target 3D objects, and physically move throughout the virtual environment (Alfadil, 2020). This context offers learners a rich multisensory, interactive experience, which is similar to the learning experience in realistic learning environments. Some studies have demonstrated that VR is an effective tool for language acquisition (e.g., Chen et al., 2020; Legault et al., 2019; Xie et al., 2019). A study by Jiao et al. (2023) asked Chinese–English bilinguals to learn two sets of L2 German words in either a picture–word association environment or an immersive VR environment. Behavioral data collected immediately after the treatment revealed a positive effect of immersive VR environment on early lexical form acquisition of novel words. Similarly, Legault et al. (2019) found that novel words learned via a VR environment are more quickly acquired than in a word–word learning environment. However, because relevant findings on the effects of learning environment have mainly focused on learning outcomes, it remains unclear as to whether a VR environment influences or interacts with sleep-dependent consolidation effects during novel word acquisition.

Present study

In the present study, we compare whether a VR environment with multisensory experiences differentially affects sleep-dependent consolidation effects of acquiring novel words in a foreign language compared with picture–word (PW) association learning. Following the same procedures of Liu and van Hell (2020), two groups of Chinese–English bilinguals received two consecutive days of learning sessions in which they learned novel Korean words. One group learned these novel words in a VR environment and the other group did so via PW learning. After the learning session on Day 2, we tested the learning effects using a four-alternative forced-choice task (4AFC) and a primed lexical-decision task. Specifically, the 4AFC task is based on the conscious processing of lexical recognition, mainly reflecting the explicit learning effects; while the primed lexical-decision task reflects automatic processing and implicit learning effects (Palma & Titone, 2021; Tartaro et al., 2021).

For both explicit and implicit learning measures, one primary concern was the role of sleep interval by comparing words learned on Day 1 (remote condition) to those learned on Day 2 (recent condition). As the CLS model indicated that off-line consolidation would contribute to novel word acquisition (Liu & van Hell, 2020; Palma & Titone, 2021), we expect that remotely learned words with a sleep interval will be better consolidated in the explicit 4AFC task as reflected by a faster recognition speeds of remotely learned words, and in the implicit primed lexical-decision task by showing faster responses to semantic relatedness decisions on remotely learned words. For the role of learning environment, given that the embodied cognition theory emphasizes the contribution of sensorimotor experiences (Barsalou, 2008), we predict better learning performance for words learned in the VR environment compared with words learned in the PW environment. More importantly, we anticipate an influence of learning environment on the sleep-dependent effect, showing a greater sleep-dependent effect (namely remotely vs. recently learned words) in the VR environment than in the PW environment. Through these comparisons, we are thus able to examine whether the sleep-dependent consolidation effect is influenced by learning environment.

Method

Participants

Sixty-seven right-handed college students (48 females, M_{age} = 20.30 years, *SD* = 1.63) were recruited and randomly divided into a VR or PW group. Seven participants were excluded because of their low accuracy (<75%) in the 4AFC task, thus leaving 60 participants to be included in the statistical analyses (VR group: $N = 31$; PW group: $N = 29$). This sample exceeded the minimum sample size of 54 calculated by G*Power 3.1 (repeated-measures analysis of variance [ANOVA], $f = .25$, $\alpha = .05$, power = .95, number of groups = 2). Participants were asked to complete a language background questionnaire in which they rated their L1 Chinese and L2 English proficiency levels based on a 7-point scale (1 = *very poor*, 7 = *excellent*). Paired-samples *t* tests revealed that the L1 proficiency ratings ($M = 6.49$, $SD = .68$) were significantly higher than L2 ratings ($M = 4.45$, $SD = .81$), *t*

= 16.26, $p < .05$. However, there was no significant difference between the two groups in terms of L1 proficiency (VR = 6.59, PW = 6.39), $t(58) = 1.14$, $p = .26$, or L2 proficiency (VR = 4.43, PW = 4.46), $t(58) = -.18$, $p = .86$. Moreover, all participants reported no prior knowledge of Korean, the third language (L3) in which the novel words were to be learned in the experiment. The study was approved by the research ethics committee at the same university, and all participants provided their informed consent before participating in the study.

On Day 2, the participants completed a survey adapted from the Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989; Lei et al., 2022) about their sleep quality over the past month and on the night of Day 1. The PSQI consists of 18 self-rated items, four of which are open-ended questions and the other 14 items are scored on a 4-point Likert scale (0–3). The PSQI provides subscale measures of seven components: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. The total score of the PSQI ranges from 0 to 21, with lower scores indicating better sleep quality. The Chinese version of PSQI has been shown to be reliable and valid when assessing the sleep quality among adults in China (Liu et al., 1996). The participants reported having good sleep quality during the last month and 92% reported having normal or higher than usual sleep quality on the night of Day 1. There was no significant difference in sleep duration on the night of Day 1 between the two groups: PW group ($M = 7.09$, $SD = .84$); VR group ($M = 7.27$, $SD = .84$), $t(58) = .83$, $p = .41$.

Materials

The materials in the present study included 30 single Korean Hangul characters, 30 2D line drawings, 30 3D objects, and a virtual scene. The Korean Hangul characters consisted of 4–9 strokes ($M = 5.93$, $SD = 1.26$). The sounds of the characters were recorded by a highly proficient Korean female speaker. The rationale for choosing Korean as the target language is because all participants reported no prior knowledge or experience with Korean. Previous work examining Chinese–English bilinguals has also selected Korean as the target language (Li et al., 2019; Qu et al., 2021).

For the VR learning environment, a virtual kitchen environment was presented and edited by the software Unity (https://unity.com). Moreover, the 3D objects were selected from a standardized database (Peeters, 2018) by considering the familiarity, visual complexity, and the appropriateness of their fit in the virtual environment (i.e., a kitchen). For the PW learning environment, all 2D line drawings corresponding to selected 3D objects were chosen from Snodgrass and Vanderwart's (1980) standardized picture list (also see Zhang & Yang, 2003). The pairings of 3D objects and 2D line drawings were randomly assigned to the Korean Hangul characters and served as semantics of the characters (Qu et al., 2021).

For the primed lexical-decision task, two Chinese semantically related prime words and two Chinese semantically unrelated prime words were selected for each Korean Hangul character. We asked a control group of 22 L1 Chinese speakers with a similar L2 background and proficiency in English as our participants to assess whether the Chinese prime words were related to the target Korean words on a 7-point scale (1 = *very unrelated*, 7 = *very related*). Results showed that the relatedness of the semantically related word pairs was significantly higher than for the semantically unrelated word pairs, $t(118) = 17.09$, $p < .001$.

Procedure

The study consisted of two learning sessions over the course of two consecutive days. After the learning session on Day 2, both groups were required to complete a 4AFC task and a primed lexical-decision task. An overview of the procedure can be seen in Fig. 1.

Learning sessions

The VR and PW groups learned the same 30 Korean Hangul characters, with 15 of the words (Set 1) being learned on Day 1 (remote words) and the other 15 words (Set 2) being learned on Day 2 (recent words). The sets of novel words presented in the two training sessions were counterbalanced across participants. A crucial difference between the remote and recent words is that the former, but not the latter, has the opportunity for sleep-dependent memory consolidation. Participants in the PW group learned novel words on a computer screen which displayed Korean Hangul characters and 2D line-drawings. By clicking the mouse, the Korean pronunciation of the character was heard. For the VR group, participants learned novel words in a simulated kitchen environment supported by VR technology. Prior to learning the Korean words, the participants were shown how to use and interact with the VR equipment. After the familiarity with the equipment, they maneuvered around and interacted with target objects in the virtual kitchen environment. They used the handset to laser point to 3D objects. Upon selecting the objects, they saw a target Korean Hangul character and heard its pronunciation. Compared with PW learning, the VR environment provides learners with a first-person, dynamic experience and an opportunity to actively explore an environment, fostering an immersive experience (Chen, 2016). On each day, both PW and VR groups received a training session for 30 minutes which consisted of 15 minutes of fixed order learning and 15 minutes of free learning.

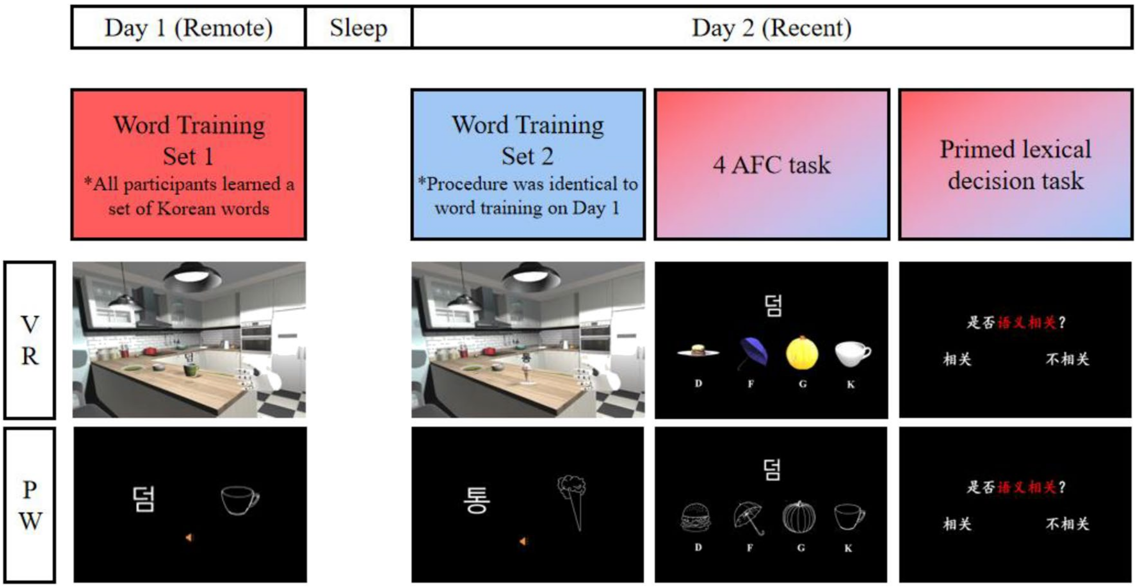

Fig. 1 Experiment procedure. The last two rows provide examples of a trial (the third row: VR environment; the fourth row: PW environment). (Color figure online)

Test session

After the word-learning session on Day 2, both groups performed a 4AFC task and a primed lexical-decision task in a fixed order. In the 4AFC task, each novel word was presented with four pictures and participants were required to select the correct picture of the target word. The pictures were 2D line drawings for the PW group, and were 2D versions of the 3D objects for the VR group. There were 30 trials consisting of 15 trials for remote words and 15 trials for recent words. Each trial began with a fixation point on a computer screen for 600 ms followed by a visually presented novel word with four pictures. Once participants gave an answer by pressing the button corresponding to one of the pictures, a blank screen was presented for 1,000 ms prior to the next trial. Before the formal experiment, the participants were presented with 5 practice trials to familiarize themselves with the procedure.

The primed lexical-decision task consisted of four blocks, with each block including 30 trials which were presented in a random order. Each participant performed 30 remote-related trials, 30 remote-unrelated trials, 30 recent-related trials, and 30 recent-unrelated trials. Each target novel word was matched with four L1 prime words (two semantically related primes and two semantically unrelated primes) and was only presented once in the same block. Trials began with a fixation cross that was presented at the center of a computer screen for 600 ms. Next, a prime word (Chinese) was presented for 250 ms, followed by a blank screen for 250 ms. A target word (Korean) was then presented for 1,000 ms and participants were required to determine whether the prime word was related to the target word or not by pressing the left or right button on the keyboard. The response keys were counterbalanced across participants. Once a response was given, a blank screen was presented for 1,000 ms prior to the next trial. Before the formal experiment, a practice block of five trials was presented to help participants become familiar with the procedure.

Results

Comprehension-based recognition task (4AFC task)

A 2 (learning environment: PW vs. VR) × 2 (condition: remote vs. recent) repeated-measures ANOVA was performed on the accuracy data. The results showed no main effects or interaction of learning environment or condition ($ps > .05$). For the analyses of reaction times (RTs), incorrect trials and trials that were ±2 standard deviations from the mean were excluded (Bakker et al., 2015). An ANOVA on RTs showed that the main effect of condition was significant, $F(1, 58) = 22.84$, $p < .001$, $\eta_p^2 = .28$, indicating faster responses to remote words ($M = 3887$, $SD = 144$) than to recent words ($M = 4920$, $SD = 148$). However, the main effect of learning environment, $F(1, 58) = .02$, $p = .89$, $\eta_p^2 < .001$, and their interaction was not significant, $F(1, 58) = 3.33$, $p = .07$, $\eta_p^2 = .05$ (see Fig. 2).

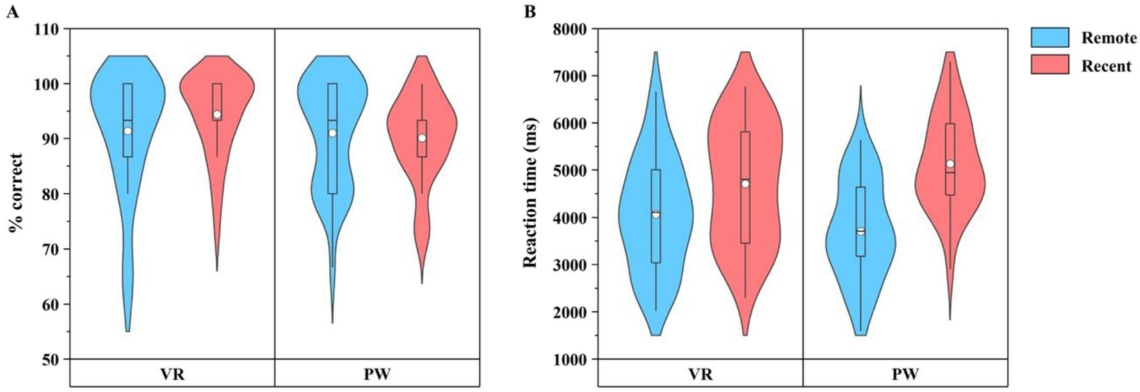

Fig. 2 Violin plots for accuracy scores (**A**) and reaction time (**B**) in the 4AFC task. The boxplot shows the interquartile range. White dots represent means. (Color figure online)

Primed lexical-decision task

For the primed-lexical decision task, we included two within-subject variables (semantic relatedness and condition) and one between-subject variable (learning environment). The ANOVA analysis on accuracy showed a significant main effect of condition, $F(1, 58) = 5.84$, $p = .02$, $\eta_p^2 = .09$, which indicated that accuracy was higher for remote words ($M = 87\%$, $SD = .99$) than for recent words ($M = 84\%$, $SD = 1.08$). The main effect of learning environment was also significant, $F(1, 58) = 4.71$, $p = .03$, $\eta_p^2 = .08$, showing a higher accuracy for VR ($M = 88\%$, $SD = 1.20$) than for PW ($M = 84\%$, $SD = 1.24$). Moreover, the main effect of semantic relatedness was significant, $F(1, 58) = 46.98$, $p < .001$, $\eta_p^2 = .45$, showing a higher accuracy for unrelated prime–target words ($M = 90\%$, $SD = 1.04$) than for related prime–target words ($M = 81\%$, $SD = 1.14$). The two-way and three-way interactions were not significant ($ps > .05$).

Incorrect trials and RTs ±2 standard deviations from the mean or over 5,000 ms were removed from the RTs analyses. The results of RTs showed a significant main effect of condition, $F(1, 58) = 24.32$, $p < .001$, $\eta_p^2 = .30$, suggesting faster responses to remote words ($M = 1017$, $SD = 36$) than to recent words ($M = 1133$, $SD = 43$). The main effect of semantic relatedness was also significant, $F(1, 58) = 158.35$, $p < .001$, $\eta_p^2 = .73$, showing faster responses to related prime–target words ($M = 897$, $SD = 31$) than unrelated prime–target words ($M = 1{,}254$, $SD = 48$). Moreover, a significant interaction between condition and semantic relatedness, $F(1, 58) = 6.36$, $p = .01$, $\eta_p^2 = .10$, further revealed that despite semantic priming effects (semantically related vs. semantically unrelated pairs) across remote words, $F(1, 59) = 77.85$, $p < .001$, $\eta_p^2 = .57$, and recent words, $F(1, 59) = 95.80$, $p < .001$, $\eta_p^2 = .62$, an enhanced semantic priming effect on recent words (related: $M = 921$, $SD = 37$; unrelated: $M = 1{,}345$, $SD = 57$) was observed compared with remote words (related: $M = 872$, $SD = 32$; unrelated: $M = 1{,}163$, $SD = 46$).

Although there was no main effect of learning environment, $F(1, 58) = .33$, $p = .57$, $\eta_p^2 = .01$, the interaction between condition and learning environment was significant, $F(1, 58) = 6.41$, $p = .01$, $\eta_p^2 = .10$. Further analyses showed that the comparison between remote words and recent words (i.e., sleep-dependent effect) was significant in the VR group (remote: $M = 1{,}010$, $SD = 50$; recent: $M = 1{,}185$, $SD = 60$) $F(1, 58) = 28.81$, $p < .001$, $\eta_p^2 = .33$, but only marginally significant in the PW group (remote: $M = 1025$, $SD = 51$; recent: $M = 1082$, $SD = 62$), $F(1, 58) = 2.79$, $p = .10$, $\eta_p^2 = .05$, suggesting that learning environment has an influence on the sleep-dependent effect. Moreover, neither the two-way interaction between learning environment and semantic relatedness, $F(1, 58) = .08$, $p = .77$, $\eta_p^2 = .001$, nor the three-way interaction, $F(1, 58) = .31$, $p = .58$, $\eta_p^2 = .005$, reached significance (see Fig. 3).

Discussion

The present study examined the influence of learning environments (PW vs. VR) on sleep-dependent consolidation effects from explicit and implicit aspects during novel word acquisition. In our analyses, we focused on two comparisons. First, the comparison between remote words after having slept and recent words without having slept was significant in both explicit and implicit measures, suggesting that a sleep-dependent consolidation effect occurred. Second, in the implicit measure, the significantly faster responses to remote words compared with recent words was only observed in the VR environment, suggesting that learning environment may play a role in sleep-dependent consolidation during novel words acquisition. We will elaborate on our main findings in the next subsections.

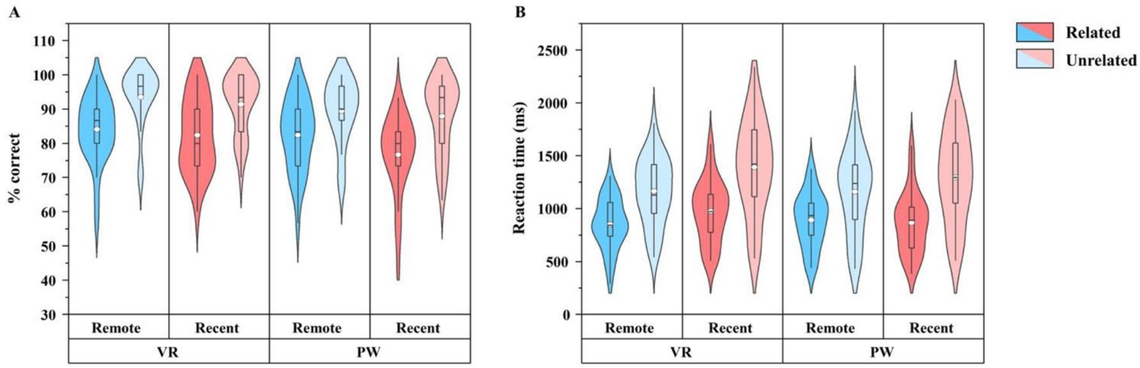

Fig. 3 Violin plots for accuracy scores (**A**) and reaction time (**B**) in the primed lexical-decision task. The boxplot shows the interquartile range. White dots represent means. (Color figure online)

Sleep-dependent consolidation

A main finding of the present study is the presence of a sleep-dependent consolidation effect which was reflected by the significant main effect of condition (remote vs. recent) in both explicit and implicit measures. On the one hand, the explicit recognition task (4AFC task) revealed faster RTs to remote words compared with recent words, consistent with some previous studies (e.g., Liu & van Hell, 2020). Although Y. Liu and van Hell adopted a recall task which differed from the recognition task in the present study, both of their measures examined sleep-dependent consolidation effects in explicit learning. The results of their recall task revealed higher accuracy for novel words learned in the remote condition than in the recent condition.

On the other hand, our findings revealed that sleep-dependent consolidation effects implicitly measured by a primed lexical-decision task also occurred, in which there was a significant difference between remote and recent words in both accuracy and RTs. This finding is in line with our expectation that remotely learned novel words receive off-line consolidation during a sleep interval, establishing better links with the existing lexical network. Moreover, the significant semantic priming effects on remote words, to some extent, also support sleep-dependent consolidation. Based on previous studies (Tamminen & Gaskell, 2013), a semantic priming effect on remote words may imply that remote words after a sleep interval are better integrated into the semantic network and might establish a more stable link with existing words. However, we unpredictably observed a semantic priming effect in recent words without sleep. This is likely because off-line consolidation may also take shape in the interval between learning and testing without sleep (Geukes et al., 2015; Lindsay & Gaskell, 2013; Szmalec et al., 2012).

The role of learning environment

The other key finding of present study is the role of learning environment observed only in the implicit measures. In the primed lexical-decision task, novel words learned via the VR environment had higher accuracy than words learned in the PW environment, which was reflected by the significant main effect of Learning environment on accuracy. Consistent with previous studies, this finding indicates that the multisensory experience supported by VR technology has a positive effect on novel word learning and supports the embodied cognition theory (Legault et al., 2019; Li & Jeong, 2020). Moreover, we observed that the learning environment interacted with sleep effects (i.e., comparison between remote and recent words) in the RT analyses, such that there was evidence supporting the larger sleep-dependent effects in the VR environment than in the PW environment.

The facilitative effects of the VR environment on sleep-dependent consolidation align with our hypotheses. According to the CLS model, in the rapid initial familiarity stage, novel words and relevant information may be quickly encoded into episodic memory and then transferred to long-term stable representations during lexical consolidation (Davis & Gaskell, 2009; McClelland et al., 1995). Compared with the unisensory input from a traditional learning environment, the multisensory experience from a VR environment may benefit the rapid encoding of novel words and relevant information into episodic memory, improving the development of episodic memory by interacting with target objects and surrounding environment (Morganti et al., 2013; Repetto et al., 2016). Consequently, the enriched episodic memory established in VR learning promotes the subsequent process of sleep-dependent consolidation, resulting in better consolidation effects via a VR environment than traditional environments. However, it was unexpected that the influence of the VR environment on sleep-dependent effects only occurred in the implicit, but not explicit learning test. One possible explanation for this finding might be related to the specificity of encoding and retrieval process in memory (Godden & Baddeley, 1975; Tulving & Thomson, 1973). The advantages of a VR environment mainly stem from rich surroundings which have been encoded implicitly, thus the helpful information may be easier to retrieve in the same implicit way as the encoding process. Future work should consider controlling for the encoding and retrieval process to further explore VR environment effect.

We acknowledge that there are some unexpected findings in the primed lexical-decision task. First, that accuracy for unrelated trials was higher than related trials, although this is not an isolated finding because previous studies about novel word acquisition have also reported similar results (e.g., Bakker et al., 2015; Kaczer et al., 2018; Liu & van Hell, 2020). It is possible that novel words may not have generated stable representations as existing words because lexical integration is a gradual and protracted process, making the semantic priming effects seem unlike existing words. Second, the RT comparisons between remote words and recent words in the PW environment were unexpected and merit further investigation. One possible explanation is that novel word learning with visually presented pictures and words may trigger superficial learning, and the limited information about novel words may not be sufficient to reflect the consolidation effect in RTs. This finding is an open question worthy of further investigation. Finally, the present study was limited to analyses of behavioral performance. Future studies may benefit from using the high temporal resolution offered by EEG technology to explore the consolidation processes elicited by a VR learning environment.

Conclusion

The present study investigated whether the rich sensory experience of a VR environment affects sleep-dependent consolidation effects during novel word learning in a foreign language as compared with a traditional learning environment. Chinese native speakers learned Korean words either in a VR environment or PW environment. For both groups, half of the words were learned on Day 1 (remote) and the other half on Day 2 (recent). Sleep-dependent consolidation effects were observed in subsequent explicit recognition and implicit lexical detection tasks, showing better performance on remote words with a sleep interval than on recent words without a sleep interval. More importantly, sleep-dependent consolidation effects were influenced by the learning environment, such that novel words learned via VR were more consolidated. Overall, the present study has revealed a facilitation effect of VR environment on sleep-dependent consolidation and offers enriching new evidence that supports the CLS model.

Author note The work was supported by the National Natural Science Foundation of China (62107024), the Humanity and Social Science Youth foundation of Ministry of Education (22YJC190015), the Natural Science Foundation of Shandong Province (ZR2021QF012), and the Institute of Psychology, CAS (GJ202004).

Declarations

Informed consent Informed consent was obtained from all individual participants included in the study.

Ethics approval This study was approved and consented by the Ethics Committee of The Department of Psychology of Qingdao University.

Conflict of interest The authors report there are no competing interests to declare.

References

Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers & Education, 153*, Article 103893. https://doi.org/10.1016/j.compedu.2020.103893

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language, 73*, 116–130. https://doi.org/10.1016/j.jml.2014.03.002

Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input processing approach. *Foreign Language Annals, 37*(2), 200–208. https://doi.org/10.1111/j.1944-9720.2004.tb02193.x

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Buysse, D. J., Reynolds, C. F., III., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Chen, J. C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education, 102*, 152–171. https://doi.org/10.1016/j.compedu.2016.08.004

Chen, Y., Smith, T. J., York, C. S., & Mayall, H. J. (2020). Google Earth virtual reality and expository writing for young English learners from a funds of knowledge perspective. *Computer Assisted Language Learning, 33*(1/2), 1–25. https://doi.org/10.1080/09588221.2018.1544151

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3773–3800. https://doi.org/10.1098/rstb.2009.0111

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35–39. https://doi.org/10.1111/j.1467-9280.2007.01845

Geukes, S., Gaskell, M. G., & Zwitserlood, P. (2015). Stroop effects from newly learned color words: Effects of memory

consolidation and episodic context. *Frontiers in Psychology, 6*, Article 278. https://doi.org/10.3389/fpsyg.2015.00278
Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*(3), 325–331.
Graves, E. A. (1936). The effect of sleep upon retention. *Journal of Experimental Psychology, 19*, 316–322. https://doi.org/10.1037/h0057827
Hacking, J. F., & Tschirner, E. (2017). The contribution of vocabulary knowledge to reading proficiency: The case of college Russian. *Foreign Language Annals, 50*(3), 500–518. https://doi.org/10.1111/flan.12282
Jiao, L., Zhu, M., Xu, Z., Zhou, G., Schwieter, W. J., & Liu, C. (2023). An ERP study on novel word learning in immersive virtual reality context. *Bilingualism: Language and Cognition*, 1–9. Advance online publication. https://doi.org/10.1017/S136672892300038X
Kaczer, L., Bavassi, L., Petroni, A., Fernández, R. S., Laurino, J., Degiorgi, S., Pedreira, M. E. (2018). Contrasting dynamics of memory consolidation for novel word forms and meanings revealed by behavioral and neurophysiological markers. *Neuropsychologia*, *117*, 472–482. https://doi.org/10.1016/j.neuropsychologia.2018.07.001
Kern, R. G. (1989). Second language reading strategy instruction: Its effects on comprehension and word inference ability. *The Modern Language Journal, 73*(2), 135–149. https://doi.org/10.2307/326569
Klinzing, J. G., Rasch, B., Born, J., & Diekelmann, S. (2016). Sleep's role in the reconsolidation of declarative memories. *Neurobiology of Learning and Memory, 136*, 166–173. https://doi.org/10.1016/j.nlm.2016.10.004
Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences, 20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004
Kurdziel, L. B., & Spencer, R. M. (2016). Consolidation of novel word learning in native English-speaking adults. *Memory, 24*(4), 471–481. https://doi.org/10.1080/09658211.2015.1019889
Kurdziel, L. B., Mantua, J., & Spencer, R. M. (2017). Novel word learning in older adults: A role for sleep? *Brain and Language, 167*, 106–113. https://doi.org/10.1016/j.bandl.2016.05.010
Lan, Y. J., Fang, S. Y., Legault, J., & Li, P. (2015). Second language acquisition of Mandarin Chinese vocabulary: Context of learning effects. *Educational Technology Research and Development, 63*(5), 671–690.
Legault, J., Fang, S. Y., Lan, Y. J., & Li, P. (2019). Structural brain changes as a function of second language vocabulary training: Effects of learning context. *Brain and Cognition, 134*, 90–102. https://doi.org/10.1016/j.bandc.2018.09.004
Lei, D., Liu, Y., & van Hell, J. G. (2022). Novel word learning with verbal definitions and images: Tracking consolidation with behavioral and event-related potential measures. *Language Learning, 72*(4), 941–979. https://doi.org/10.1111/lang.12502
Li, P., & Jeong, H. (2020). The social brain of language: Grounding second language learning in social interaction. *npj Science of Learning, 5*(1), 1–9. https://doi.org/10.1038/s41539-020-0068-7
Li, H., Qu, J., Chen, C., Chen, Y., Xue, G., Zhang, L., ... Mei, L. (2019). Lexical learning in a new language leads to neural pattern similarity with word reading in native language. *Human Brain Mapping*, *40*(1), 98–109. https://doi.org/10.1002/hbm.24357
Lindsay, S., & Gaskell, M. G. (2013). Lexical integration of novel words without sleep. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 608–622. https://doi.org/10.1037/a0029243
Liu, Y., & van Hell, J. G. (2020). Learning novel word meanings: An ERP study on lexical consolidation in monolingual, inexperienced foreign language learners. *Language Learning, 70*(S2), 45–74. https://doi.org/10.1111/lang.12403
Liu, X., Tang, M., Hu, L., Wang, A., Wu, H., Zhao, G., Gao, C., & Li, W. (1996). The reliability and validity of the Pittsburgh Sleep Quality Index. *Chinese Journal of Psychiatry, 29*(2), 103–107. [In Chinese].
McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227–234. https://doi.org/10.1037/h0031564
Morganti, F., Stefanini, S., & Riva, G. (2013). From allo-to egocentric spatial ability in early Alzheimer's disease: A study with virtual reality spatial tasks. *Cognitive Neuroscience, 4*(3/4), 171–180. https://doi.org/10.1080/17588928.2013.854762
Palma, P., & Titone, D. (2021). Something old, something new: A review of the literature on sleep-dependent memory consolidation of novel words in adults. *Psychonomic Bulletin & Review, 28*(1), 96–121. https://doi.org/10.3758/s13423-020-01809-5
Peeters, D. (2018). A standardized set of 3-D objects for virtual reality research and applications. *Behavior Research Methods, 50*(3), 1047–1054. https://doi.org/10.3758/s13428-017-0925-3
Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246–256. https://doi.org/10.1037/0022-0663.97.2.246
Qu, J., Hu, L., Liu, X., Dong, J., Yang, R., & Mei, L. (2021). The contributions of the left hippocampus and bilateral inferior parietal lobule to form-meaning associative learning. *Psychophysiology, 58*(8), Article e13834. https://doi.org/10.1111/psyp.13834
Repetto, C., Serino, S., Macedonia, M., & Riva, G. (2016). Virtual reality as an embodied tool to enhance episodic memory in elderly. *Frontiers in Psychology, 7*, Article 1839. https://doi.org/10.3389/fpsyg.2016.01839
Smidt, E., & Hegelheimer, V. (2004). Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition, and strategy use. *Computer Assisted Language Learning, 17*(5), 517–556. https://doi.org/10.1080/0958822042000319692
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174
Szmalec, A., Page, M. P., & Duyck, W. (2012). The development of long-term lexical representations through Hebb repetition learning. *Journal of Memory and Language, 67*(3), 342–354. https://doi.org/10.1016/j.jml.2012.07.001
Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *Quarterly Journal of Experimental Psychology, 66*(5), 1001–1025. https://doi.org/10.1080/17470218.2012.7246
Tartaro, G., Takashima, A., & McQueen, J. M. (2021). Consolidation as a mechanism for word learning in sequential bilinguals. *Bilingualism: Language and Cognition, 24*(5), 864–878. https://doi.org/10.1017/S1366728921000286
Tham, E. K., Lindsay, S., & Gaskell, M. G. (2015). Markers of automaticity in sleep-associated consolidation of novel words. *Neuropsychologia, 71*, 146–157. https://doi.org/10.1016/j.neuropsychologia.2015.03.025
Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Wang, H. C., Savage, G., Gaskell, M. G., Paulin, T., Robidoux, S., & Castles, A. (2017). Bedding down new words: Sleep promotes the emergence of lexical competition in visual word recognition. *Psychonomic Bulletin & Review, 24*(4), 1186–1193. https://doi.org/10.3758/s13423-016-1182-7

Xie, Y., Chen, Y., & Ryder, L. H. (2019). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning, 34*(3), 225–245. https://doi.org/10.1080/09588221.2019.1604551

Zhang, Q., & Yang, Y. (2003). The determiners of picture-naming latency. *Acta Psychologica Sinica, 35*(04), 447.

Open practices statement The data are available from the corresponding author on reasonable request.