

Neural tracking of newly learned word lexicalisation in virtual reality: the effects of emotional prosody

Yue Lin, John W. Schwieter, Lu Jiao & Cong Liu

To cite this article: Yue Lin, John W. Schwieter, Lu Jiao & Cong Liu (26 Mar 2026): Neural tracking of newly learned word lexicalisation in virtual reality: the effects of emotional prosody, Language, Cognition and Neuroscience, DOI: [10.1080/23273798.2026.2650100](https://doi.org/10.1080/23273798.2026.2650100)

To link to this article: <https://doi.org/10.1080/23273798.2026.2650100>

 [View supplementary material](#) 

 Published online: 26 Mar 2026.

 [Submit your article to this journal](#) 

 Article views: 4

 [View related articles](#) 

 [View Crossmark data](#) 

Neural tracking of newly learned word lexicalisation in virtual reality: the effects of emotional prosody

Yue Lin^{a,b}, John W. Schwieter^{d,e,f}, Lu Jiao^{a,c} and Cong Liu^{a,c}

^aDepartment of Psychology, School of Education Science, Qingdao University, Qingdao, People's Republic of China; ^bSchool of Psychology, South China Normal University, Guangzhou, People's Republic of China; ^cBrain, Cognition, and Language Learning Laboratory, Qingdao University, Qingdao, People's Republic of China; ^dLanguage Acquisition, Multilingualism, and Cognition Laboratory/Bilingualism Matters, Wilfrid Laurier University, Waterloo, Canada; ^eMcMaster University, Hamilton, Canada; ^fSchool of Education, University College Dublin, Dublin, Ireland

ABSTRACT

In this behavioural and electrophysiological study, we examined the effects of emotional prosody on lexicalisation of newly-learned words in an immersive virtual reality (VR) environment. Chinese speakers learned German words that were accompanied by positive, negative, or neutral emotional prosody. Lexicalisation was subsequently assessed in a semantic priming task while brain activity was simultaneously measured through EEG. The results showed that emotional prosody, regardless of valence, slowed responses to newly learned words relative to those learned with neutral prosody. Critically, words learned under positive prosody conditions elicited larger semantic priming effects in both behavioural performance and the late positive component (LPC). Moreover, the magnitude of this effect was significantly correlated with the participants' self-reported emotional experience during VR learning. Taken together, our findings suggest that in a VR learning environment, positive prosody facilitates lexicalisation of newly learned words, providing empirical support for the Integrated Cognitive-Affective Theory of Learning with Media.

ARTICLE HISTORY

Received 25 April 2025
Accepted 8 March 2026

KEYWORDS

Lexicalisation; emotional prosody; virtual reality; EEG

1. Introduction

An important question in language sciences is how language learning and word retrieval is affected by extra-linguistic factors such as emotional prosody. In the literature on second language acquisition, research over the past few decades has systematically investigated the effects of emotions on word learning and lexicalisation in traditional learning environments (Chiu & Churchill, 2016; Fyfield et al., 2022; Guo et al., 2018). Recently, given the rapid development of technology, researchers in language learning have increasingly examined these issues in immersive environments, such as virtual reality (VR) (Jiao et al., 2024; Legault et al., 2019; Liu et al., 2024). However, whether emotions play a modulating role in VR-assisted language learning has not yet been examined. In daily conversations, speakers not only use words to convey what they want to communicate, but they also use other elements, both verbal and nonverbal. Emotional prosody, one of the most efficient modalities for expressing emotion (Jungers et al., 2024; Schirmer & Kotz, 2003), is important for spoken word recognition (Sumner et al., 2014).

Examining these issues can help to reveal the potential contributions of emotional cues in VR-assisted language learning.

1.1. Effects of emotional cues

The *Integrated Cognitive-Affective Theory of Learning with Media* (ICALM; Plass & Kaplan, 2016) posits that emotional cues modulate learners' cognitive processing, particularly the selection and organisation of information in working memory, thereby shaping learning outcomes (see also Moreno & Mayer, 2007). In the language learning literature, for example, Guo et al. (2018) asked Chinese speakers to learn words using picture-word associations that were accompanied by positive, neutral, or negative visual cues (i.e. emotional pictures). Their performance on a subsequent semantic category judgement task revealed that negative emotional cues showed an overall suppressing effect on learning new words, whereas positive emotional cues facilitated word learning only in a shallower semantic category learning task.

CONTACT Cong Liu  congu902@gmail.com  Department of Psychology, Qingdao University, No.308, NingXia R., Qingdao 266071, People's Republic of China

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23273798.2026.2650100>.

© 2026 Informa UK Limited, trading as Taylor & Francis Group

In addition to visual emotional cues, auditory emotional cues such as emotional prosody can effectively convey emotions, feelings, and attitudes through systematic modulation of tone (Mauchand & Pell, 2021; Schirmer & Kotz, 2003). As a form of sound-based social information, emotional prosody serves more global communicative functions, such as signalling a speaker's overall affective state (Kim & Sumner, 2017). For example, Brase and Mani (2017) examined the effects of emotional prosody on word learning in a video-based learning environment among German–English bilinguals. Learners were presented with videos in which an actor explained the meaning of each new word in a sentence, delivered either with emotional prosody or neutral prosody. In the emotional condition, explanations were produced with an expressive tone of voice accompanied by gestures, whereas in the neutral condition, explanations were delivered in a neutral tone without gestures. The results showed that emotional prosody facilitated word learning in a foreign language.

However, existing research about emotional effects remains largely confined to immediate learning performance across tasks such as lexical decisions (Zhang et al., 2020), free recall (Brase & Mani, 2017), and lexical-semantic judgements (Guo et al., 2018), ignoring the effect of emotional prosody on subsequent lexicalisation of newly-learned words. According to the Complementary Learning Systems model, word learning encompasses an initial rapid familiarity stage followed by a slow lexicalisation stage (Blakeman & Mareschal, 2020; Davis & Gaskell, 2009; McClelland et al., 1995). Unlike the rapid familiarity phase, offline lexicalisation primarily reflects the gradual integration of a new word into the existing lexical network, often assessed via semantic priming tasks (Lei et al., 2022; Liu & van Hell, 2020). Accordingly, the present study aims to address the gap in understanding how emotional prosody influences this critical lexicalisation process.

1.2. Effects of virtual reality environment

According to the *Cognitive Affective Model of Immersive Learning* (CAMIL), technological features of immersive VR elicit heightened feelings of presence (i.e. the psychological sense of “being there”) and agency (i.e. being in control of one's actions) in learners, which subsequently influence learning through affective and cognitive factors such as situational interest, intrinsic motivation, and emotion (Makransky & Petersen, 2021). Within the model's framework, research on word learning has increasingly shifted from images/videos-based learning to immersive VR learning in recent years (Fuhrman et al., 2021; Li & Jeong, 2020; Liu et al., 2026). Compared

to video-based learning, VR offers a higher degree of immersion by placing learners in interactive environments that closely resemble real-life situations, allowing them to physically navigate the environment and manipulate objects (e.g. pointing and selecting). This sensorimotor experience can enhance learners' motivation and induce positive emotional experiences, thereby promoting deeper learning (Mayer, 2021).

A growing body of evidence has demonstrated that VR can be an effective tool for word learning and lexicalisation (for a review, see Chen et al., 2022). For example, Jiao et al. (2024) asked Chinese speakers to learn two sets of German words separately in an immersive VR or picture-word (PW) association environment for three days. The findings from a lexical decision task demonstrated a clear advantage of VR for early lexical-form learning. Furthermore, another more recent study by Jiao et al. (2026) found a facilitative effect of VR on later stages of lexicalisation of newly-learned words through a semantic priming task.

In sum, the research discussed above demonstrates the effects of extra-linguistic factors such as emotion and VR technology on language learning. However, less is known about how the presence of emotional prosody in an immersive VR environment benefits word learning, especially during the process of lexicalisation. A more nuanced understanding of such potential emotional effects is important as it holds implications for understanding sociolinguistic factors impacting language learning.

1.3. The present study

The present study examines how emotional prosody influences lexicalisation of newly learned words by combining behavioural measures of learning performance with neurocognitive EEG recordings. Chinese speakers were instructed to learn German words on two days in a VR environment which was accompanied by positive, neutral, and negative prosody. Subsequently, participants completed a semantic priming task to assess the lexicalisation of the newly learned words. Lexicalisation was indexed by semantic priming effects which are differences in reaction times (RTs) and accuracy between related and unrelated prime-target pairs (Bakker et al., 2015; Jiao et al., 2026; Liu & van Hell, 2020). During the task, participants' electrophysiological activity was measured using EEG. Based on previous work and our research objectives, we focused on two ERP components: N400 and late positive component (LPC), both of which have been associated with lexicalisation in studies on word learning (Lei et al., 2022; Liu & van Hell, 2020). Specifically, during lexicalisation, N400

effects are reflective of automatic processes, whereas LPC effects are associated with controlled, strategic processes. Finally, the participants completed a questionnaire assessing their emotional experience during the VR learning process.

Based on the ICALM and previous studies examining emotional effects on word learning (Brase & Mani, 2017; Guo et al., 2018), we hypothesise that positive prosody will facilitate the lexicalisation of newly learned words, whereas negative prosody will impair it. We expect that these effects will be reflected in the magnitude of semantic priming in both behavioural performance and electrophysiological activity (N400 and LPC), compared to conditions of neutral prosody. We further predict that the degree of lexicalisation of newly-learned words will be modulated by learners' subjective emotional experiences in the VR learning environment.

2. Method

2.1. Participants

Thirty-one Chinese-English bilinguals (Age = $20.06_M \pm 1.41_{SD}$) participated in the present study. This sample exceeds the minimum size of 19 calculated by G*Power 3.1: $f = .25$, $\alpha = .05$, power = .80, number of groups = 1, number of measurements = 6. All participants were first language (L1) Chinese speakers who had learned L2 English in classroom settings, with no immigration or immersion experience in a non-Chinese-speaking environment. They also had no prior knowledge of German, the language in which they learned new words. The language background questionnaire elicited participants' self-ratings of language proficiency on 7-point scales (1 = very poor, 7 = excellent). Analyses showed that participants rated their L1 Chinese significantly more proficient than English (Chinese = $6.22_M \pm .81_{SD}$, English = $3.71_M \pm .95_{SD}$, $t = 12.12$, $p < .01$). All participants were right-handed adults, with normal or corrected-to-normal vision, normal hearing, and no history of neurological or language-related disorders. The local ethics committee approved the study and all participants provided their written informed consent prior to participating in the experiment. They received a modest payment for their participation.

2.2. Materials

2.2.1. Linguistic stimuli

Twenty-four German words (see Supplementary Table 1) served as linguistic stimuli in the learning and testing sessions. The rationale for choosing German words as the to-be-learned words was because participants

reported no prior knowledge or experience with German, and some relevant studies have also selected German as target language (Jiao et al., 2021). In our study, German words were recorded by a highly-proficient Chinese-German male speaker in a soundproof room, avoiding the confounding effects of external noise. Using the Chinese-German speaker's voice helped to ensure that emotional prosody was delivered and encoded in a manner that was culturally congruent with the Chinese participants, thereby minimising potential cultural barriers in emotion recognition (Elfenbein & Ambady, 2002) while maintaining native-like pronunciation in German. Emotional prosody was presented through the tone of voice, with eight German words being presented with positive prosody, eight with negative prosody, and eight without emotional prosody to serve as a control condition. To assess emotional prosodic valence of the target German words, a group of 22 individuals who did not take part in the formal experiment were recruited. These individuals rated the valence of emotional prosody on a 7-point scale (1 = the most negative, 7 = the most positive). Analyses showed that these individuals rated the emotional valence significantly different for each condition (positive = $5.59_M \pm .28_{SD}$ > neutral = $3.52_M \pm .19_{SD}$ > negative = $2.10_M \pm .21_{SD}$).

2.2.2. Virtual reality environment and stimuli

An immersive virtual environment of a kitchen was created using Unity software (<https://unity.com>). All coloured three-dimensional (3D) objects corresponding to 24 German words were selected from a standardised database (Peeters, 2018) and presented in the virtual environment (e.g. *Kuchen*, "cake"; *Apfel*, "apple"). Using HTC VIVE headgear and a controller to laser point at target stimuli, participants fully immersed themselves in the environment, interacting with 3D objects, and hearing their pronunciation of corresponding German words. They were instructed to listen attentively to the German words to learn their meanings. No vocal repetition or production of the words was required during the learning sessions.

2.3. Procedure and measures

2.3.1. Learning sessions

Figure 1 presents an overview of the learning and testing stages. During the learning sessions, participants learned 24 German words (8 words with positive prosody, 8 words with negative prosody, and 8 words with neutral prosody). Before beginning the learning session, participants completed a language background questionnaire. They were then shown how to use the VR

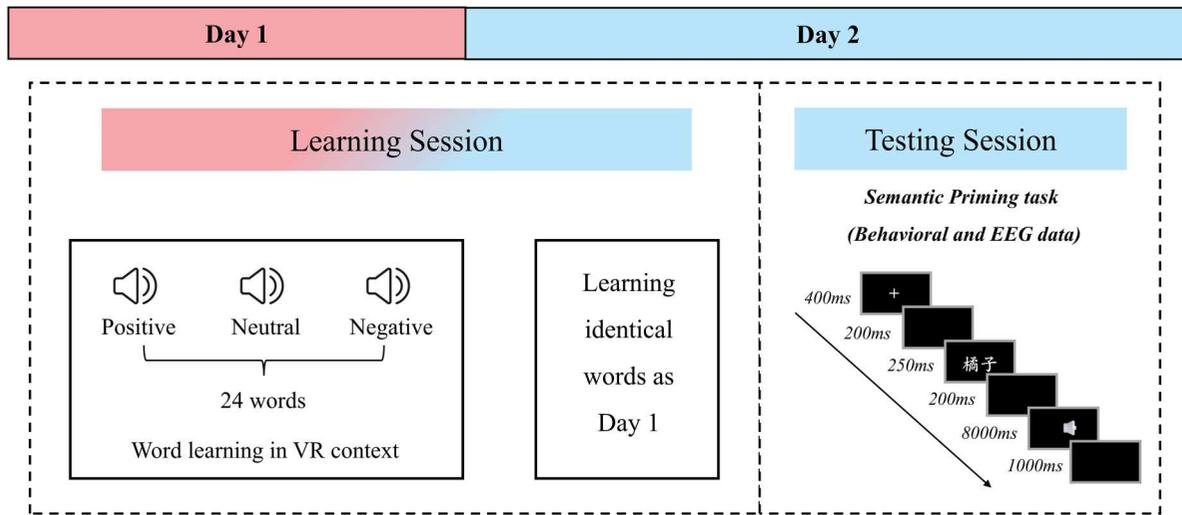


Figure 1. Schematic overview of the present study. On Day 1 and Day 2, participants learned three sets of German words through VR. A semantic priming task was completed after the learning session on Day 2.

equipment and were asked to familiarise themselves with it and the virtual environment. During this period, participants wore the headgear to receive high quality visual fidelity and were able to physically move throughout the virtual kitchen environment. Additionally, they were required to visually search for and find the target 3D objects. Upon locating a target, they used the controller to laser point and select it, hearing the name of the object in Chinese. This process ensured that participants were familiar with the identity and spatial location of all target objects prior to being exposed to them in the German learning sessions, thereby reducing visual search demands during learning.

After participants were familiar with the virtual environment and target objects, they were asked to learn all German words using the headgear and headsets. As in the practice period, participants moved throughout the VR learning environment during the learning session and selected target 3D objects to hear the corresponding German words through headphones. All target words were divided into three sets reflecting the three different conditions of emotional prosody, which was counterbalanced across the sets for different participants (e.g. participant A learned positive set 1, neutral set 2, and negative set 3; while participant B learned neutral set 1, negative set 2, and positive set 3). The three sets were presented in blocks and their order was also counterbalanced across participants. In addition, participants were able to select and hear the words as many times as they chose within a seven-minute time limit for each emotional prosodic condition. The participants learned the same sets of words for 2 days and the learning duration for these new words was about 21 min each day.

After the learning session, participants completed a questionnaire assessing their emotional experience during the immersive VR learning process. The subjective emotion questionnaire consisted of five items based on Parong and Mayer (2018), e.g. “*I felt happy during the lesson*”.

2.3.2. Testing session

After Day 2’s learning session, participants were asked to complete a semantic priming task in which behavioural and EEG data were recorded. Each newly-learned word served as a target word, and was paired with four prime words in Chinese: two semantically related and two unrelated to the target word. Crucially, to ensure consistency in lexical properties, the unrelated primes were not random words; rather, they were selected from the list of related primes for other target words and re-paired with the current target (i.e. the unrelated primes were potentially related to other targets). This design was employed to secure enough trials for ERP averaging while minimising repetition effects and strategic prediction strategies associated with fixed prime-target pairs. This procedure has been well established in the literature and implemented in numerous studies (e.g. Lei et al., 2022; Liu & van Hell, 2020). There were four blocks totalling 192 prime-target trials. Each of the 24 newly learned words appeared twice per block, once with a related prime and once with an unrelated prime, yielding 48 trials per block. Across the four blocks, each of the three emotion conditions (positive, negative, neutral) included 64 trials (16 trials \times 4 blocks), with 32 related and 32 unrelated trials. This trial count (i.e. 32 trials) provides an adequate signal-to-noise ratio for reliable ERP analysis (Boudewyn et al.,

2018). Moreover, the newly learned words in the testing session were presented with the same emotional prosody as in the learning session.

Each trial started with a 400 ms fixation cross followed by a blank screen for 200 ms. A prime word was then presented for 250 ms, followed by a blank screen for 200 ms, and the auditory presentation of the target. Participants were required to identify whether the prime and target words were semantically related or not by pressing the “F” or “J” key. The response keys were counterbalanced across participants. Once the participant pressed the response key or if no response was made within 8000 ms, a blank screen was presented for 1000 ms before the next trial began. Participants also took a brief break between each block. Before the formal test, they were given a practice block of 10 trials followed by feedback on whether they were correct or not.

2.4. EEG recordings and preprocessing

An elastic cap (Brain Products, Germany) with 64 active Ag/AgCl electrodes was placed on participants’ scalps according to the extended 10–20 positioning system. Impedances were kept below 5 k Ω . EEG signals were filtered online with a band pass of .05–100 Hz and re-filtered offline with a band pass between .5 and 30 Hz. The signal was recorded at a 1 kHz sampling rate. Electrodes were referenced to a vertex reference (electrode FCz) and re-referenced offline to an average of bilateral mastoids. Ocular artifact reduction was performed through Independent Component Analysis using EEGLAB in MATLAB software (Delorme & Makeig, 2004). Then, the continuous signals were cut into epochs ranging from 200 ms prior to stimulus onset (pre-stimulus baseline) to 1000 ms after target stimuli onset. The epochs with voltages exceeding $\pm 80 \mu\text{V}$ were automatically discarded. Moreover, we employed rigorous artifact-removal procedures to ensure the quality of the EEG data. On average, 30 trials (approximately 15%) were removed per participant. During ICA decomposition and artifact correction, we followed standard procedures to identify and exclude components associated with ocular artifacts (e.g. blinks and horizontal eye movements) (Jung et al., 2000). None of the electrodes included in the target analyses were interpolated.

2.5. Data analyses

All data were analysed with multilevel Bayesian regression models in R (R Core Team, 2020) using the *brms* package (Bürkner, 2017, 2021). Bayesian modelling has advantages over frequentist approaches such as

analysis of variance and linear mixed-effect models (Van de Schoot & Depaoli, 2014). First, Bayesian modelling presents the whole posterior probability distribution of each effect rather than reporting only a binary test for the null hypothesis. It allows us to calculate *credibility intervals* (instead of confidence intervals from a frequentist approach) and *evidence ratios* in favour of a directional hypothesis (e.g. an effect being greater than zero). Since Bayesian models do not report *p*-values, we use a hypothesis-testing method to calculate evidence ratios and credibility intervals to quantify the likelihood of a given hypothesis with respect to the alternative. Evidence ratios above 19 are roughly equivalent to “significant” at an $\alpha = .05$ level for a given hypothesis (Escudero et al., 2020). Meanwhile, if the 95% credibility interval does not contain zero (i.e. less than 2.5% of the posterior distribution located on the other side of zero), a given hypothesis can be considered “significant” at an $\alpha = .05$ level (Milne & Herff, 2020). In addition to evidence ratios and 95% credibility intervals, we also report the posterior probabilities of a given hypothesis, which may lie between 0 and 1.

Second, similar to a classical mixed effects model, Bayesian modelling contains population-level effects and group-level effects which are separately parallel to the fixed effects and random effects of the *lme4* statistical package in R. Our Bayesian models use a maximal group-level (random) effects structure with two such “groups”, namely participants and items, hence the population-level effects reflect the estimated means of effects across all participants and items. These maximal random effects structures often fail to converge using frequentist analyses but are feasible with Bayesian regressions. Population-level (fixed) effects include emotion and relatedness variables. We turned the emotion variable into dummy-coded variable (referenced to the neutral condition) and the relatedness variable into sum-coded variable (referenced to the unrelated condition).

Finally, Bayesian modelling approach incorporates prior knowledge, which is referred to as *priors*. Largely in line with previous relevant studies, we used weakly informative priors that constrained the models to psycho-linguistically plausible parameter estimates (Gelman & Carpenter, 2020; Lago et al., 2023; Murakami & Ellis, 2022). The prior for fixed effects followed a normal distribution of $N(0, 10)$. We derived the posterior distributions based on Hamiltonian Monte Carlo with four Markov chains with 5000 iterations each, including 1000 warmup iterations. All parameters had a Gelman-Rubin statistic *Rhat* equal to 1.00. Supplementary Figure 1–4 in Supporting Information showed the density plots and trace plots for the intercept and fixed-effects parameters.

RTs were entered into the Bayesian model and excluded incorrect responses (13.4%) as well as RTs \pm 2SD from the mean for each trial type (7.6%) (Liu et al., 2024). For the ERP data, given that N400 (400–600 ms; C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2) and LPC (600–750 ms; F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) were the main foci, we examined the mean amplitude of the waveforms across the selected time-window of each component in the semantic priming task. Moreover, P200 (190–260 ms; F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) was also analysed based on the grand-average waveforms obtained from the semantic priming task.

3. Results

3.1. Behavioural results

Figure 2 shows the mean RTs of the semantic priming task. We ran a multi-level Bayesian regression model to test the effect of emotional prosody and relatedness on RTs (see Table 1, see also Supplementary Figure 5 in the Supporting Information for a visualisation of the model's posterior distribution). All types of newly learned words demonstrated significant semantic priming effects, with faster RTs to related prime-target pairs than to unrelated pairs (neutral prosody: related = 1782 ms, unrelated = 1950 ms, $ER = 55.54$; positive prosody: related = 1832 ms, unrelated = 2114 ms, $ER = 5332.33$; negative prosody: related = 1909 ms, unrelated = 2124 ms, $ER = 614.38$), corresponding to very strong evidence. More importantly, the semantic priming effects in positive conditions were larger than in neutral conditions, as shown by the interaction between relatedness \times emotion (neutral vs. positive), with very strong evidence ($ER = 201.53$) and a posterior probability of 100%. Moreover, RTs to words learned with emotional prosody (both positive and negative)

were slower than those learned with neutral prosody, with very strong evidence in both comparisons ($ER = 147.15$ and $ER = 339.43$, respectively).

3.2. ERPs results

3.2.1. P200 component

Figure 3 illustrates the grand-average ERP waveforms elicited during the semantic priming task for neutral, positive, and negative prosody conditions. As shown in Table 2 and Supplementary Figure 6, only words learned with neutral prosody elicited larger P200 amplitude in unrelated pairs compared to related pairs, corresponding to strong evidence ($ER = 20.45$). This relatedness effect was absent for words learned with positive ($ER = 7.29$) or negative ($ER = 4.70$) prosody. A significant interaction between relatedness \times emotion showed a larger P200 amplitude difference between related and unrelated pairs for words learned with neutral prosody than for those learned with positive ($ER = 56.14$) or negative ($ER = 30.81$) prosody. In addition, words learned with either positive or negative prosody elicited larger P200 amplitudes than those learned with neutral prosody, with strong evidence in both comparisons ($ER = 28.47$ and $ER = 3199.00$, respectively).

3.2.2. N400 component

The Bayesian analyses found no significant main effects or interactions in N400 amplitude ($ERs < 19$, see Table 3 and Supplementary Figure 7).

3.2.3. LPC component

As shown in Table 4 and Supplementary Figure 8, a significant LPC semantic priming effect was observed only for words learned with positive prosody, with larger LPC amplitudes for unrelated than for related prime-target

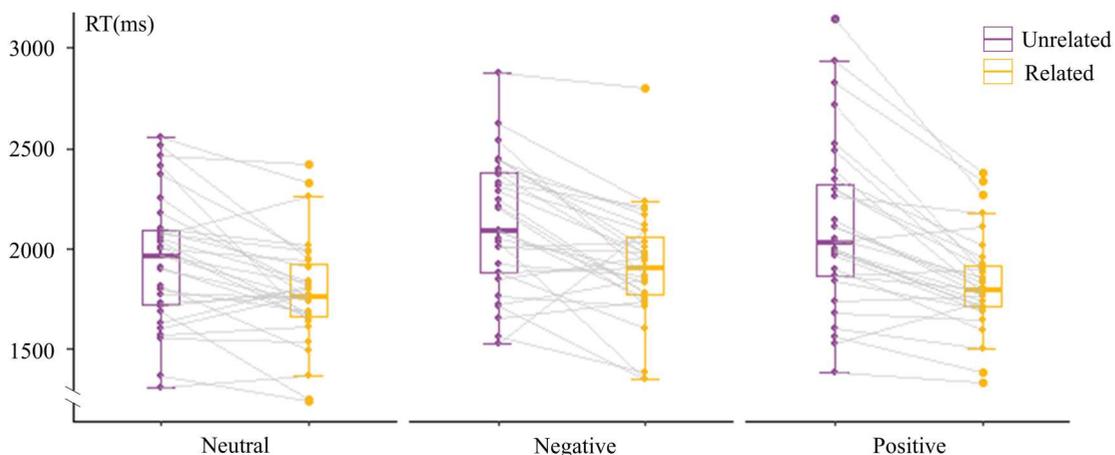


Figure 2. Box plots showing the RTs of the semantic priming task for words learned under neutral, negative, and positive conditions.

Table 1. Summary of the directional hypotheses in the RT model.

Hypothesis	Mean	SE	95% credibility interval		Evidence ratio	Posterior probability	Star
Neutral condition as the reference level							
Relatedness < 0	-.19	.09	-.34	-.04	55.54	.98	*
Emotion (neutral vs. positive) > 0	.13	.05	.05	.22	147.15	.99	*
Emotion (neutral vs. negative) > 0	.21	.07	.09	.33	339.43	1.00	*
Relatedness × Emotion (neutral vs. positive) < 0	-.18	.07	-.30	-.07	201.53	1.00	*
Relatedness × Emotion (neutral vs. negative) < 0	-.08	.09	-.23	.08	3.97	.80	
Positive condition as the reference level							
Relatedness < 0	-.37	.09	-.52	-.22	5332.33	1.00	*
Emotion (positive vs. negative) > 0	.08	.07	-.03	.19	6.99	.87	
Relatedness × Emotion (positive vs. negative) > 0	.10	.09	-.04	.25	7.19	.88	
Negative condition as the reference level							
Relatedness < 0	-.28	.09	-.42	-.13	614.38	1.00	*

Note: A star indicates if 0 lies outside the 95% CI.

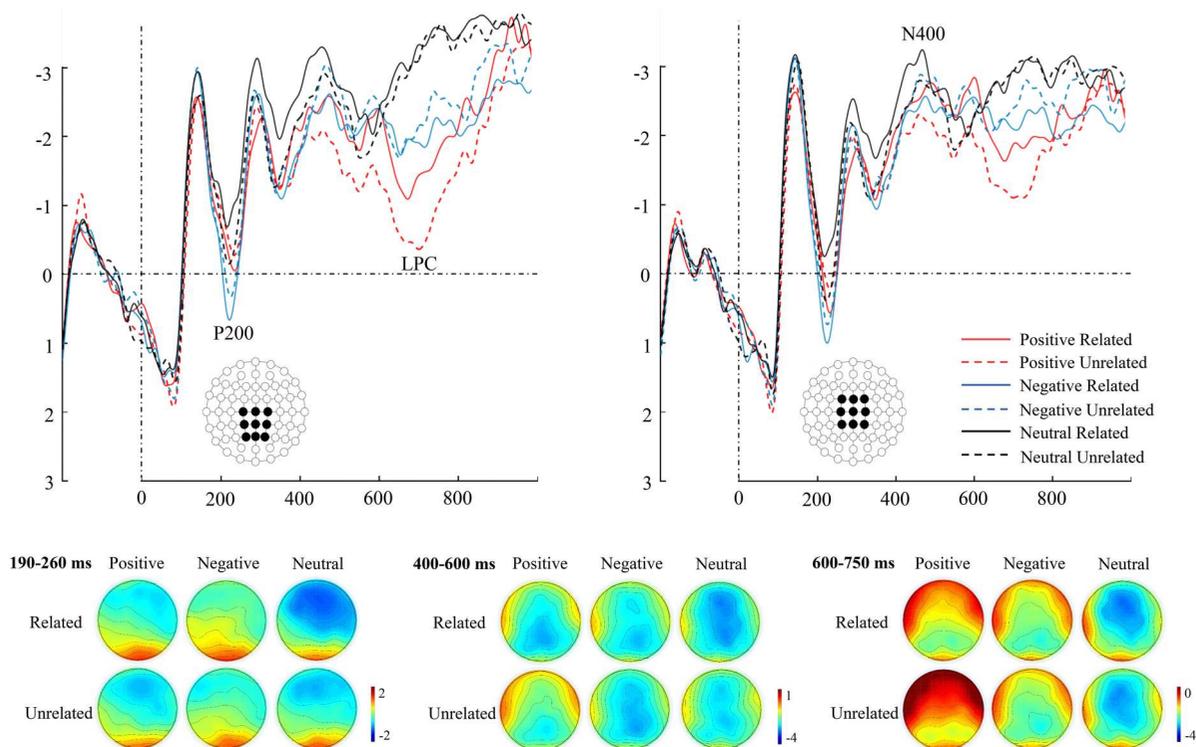


Figure 3. Grand average waveforms and topographic maps of the semantic priming task for words learned under neutral, negative, and positive conditions.

pairs ($ER = 44.85$). Critically, this LPC priming effect was significantly larger for words learned with positive prosody than for those learned with negative prosody, as indicated by an interaction between emotion × relatedness (positive vs. negative; $ER = 79.81$). In addition, words learned with either positive or negative prosody elicited larger LPC amplitudes than those learned with neutral prosody, with posterior probabilities of 100% in both comparisons.

3.3. Correlation results

To examine the relationship between the lexicalisation of newly-learned words and learners' subjective

emotional experience in the VR-assisted learning environment, Pearson correlation analyses were conducted between semantic priming effects (calculated by subtracting related trials from unrelated trials) and self-reported emotion scores. The results revealed a significant positive correlation between emotion ratings and the overall semantic priming effect, averaged across the three prosody conditions ($r = .39$, $p = .03$; see Figure 4(A)). We also conducted Pearson correlation analyses separately for each prosody condition. A significant positive correlation was observed in the positive prosody condition ($r = .41$, $p = .02$; see Figure 4(B)). In contrast, no significant correlations

Table 2. Summary of the directional hypotheses in the P200 model.

Hypothesis	Mean	SE	95% credibility interval		Evidence ratio	Posterior probability	Star
Neutral condition as the reference level							
Relatedness < 0	-.73	.45	-1.47	-.01	20.45	.95	*
Emotion (neutral vs. positive) > 0	.57	.32	.05	1.09	28.47	.97	*
Emotion (neutral vs. negative) > 0	1.13	.32	.60	1.66	3199.00	1.00	*
Relatedness × Emotion (neutral vs. positive) > 0	1.24	.59	.27	2.21	56.14	.98	*
Relatedness × Emotion (neutral vs. negative) > 0	1.14	.61	.13	2.15	30.81	.97	*
Positive condition as the reference level							
Relatedness > 0	.50	.43	-.21	1.20	7.29	.88	
Emotion (positive vs. negative) > 0	.55	.33	.01	1.09	20.00	.95	*
Relatedness × Emotion (positive vs. negative) < 0	-.11	.60	-1.10	.88	1.37	.58	
Negative condition as the reference level							
Relatedness < 0	.40	.43	-.31	1.10	4.70	.82	

Note: A star indicates if 0 lies outside the 95% CI.

Table 3. Summary of the directional hypotheses in the N400 model.

Hypothesis	Mean	SE	95% credibility interval		Evidence ratio	Posterior probability	Star
Neutral condition as the reference level							
Relatedness < 0	-.31	.35	-.89	.26	4.51	.82	
Emotion (neutral vs. positive) > 0	.47	.29	-.01	.94	17.50	.95	
Emotion (neutral vs. negative) > 0	.16	.29	-.32	.63	2.52	.72	
Relatedness × Emotion (neutral vs. positive) < 0	-.05	.51	-.89	.79	1.14	.53	
Relatedness × Emotion (neutral vs. negative) > 0	.77	.49	-.03	1.56	16.51	.94	
Positive condition as the reference level							
Relatedness < 0	-.36	.35	-.94	.22	5.33	.84	
Emotion (positive vs. negative) < 0	-.33	.30	-.82	.16	6.42	.87	
Relatedness × Emotion (positive vs. negative) > 0	.80	.49	-.01	1.60	18.42	.95	
Negative condition as the reference level							
Relatedness < 0	.46	.35	-.12	1.03	9.71	.91	

Note: A star indicates if 0 lies outside the 95% CI.

Table 4. Summary of the directional hypotheses in the LPC model.

Hypothesis	Mean	SE	95% credibility interval		Evidence ratio	Posterior probability	Star
Neutral condition as the reference level							
Relatedness < 0	-.10	.47	-.86	.66	1.40	.58	
Emotion (neutral vs. positive) > 0	3.03	.44	2.31	3.77	>9999.00	1.00	*
Emotion (neutral vs. negative) > 0	1.65	.34	1.09	2.22	>9999.00	1.00	*
Relatedness × Emotion (neutral vs. positive) < 0	-.81	.62	-1.83	.20	9.40	.90	
Relatedness × Emotion (neutral vs. negative) > 0	.62	.64	-.45	1.66	4.99	.83	
Positive condition as the reference level							
Relatedness < 0	-.91	.45	-1.64	-.18	44.85	.98	*
Emotion (positive vs. negative) < 0	-1.38	.42	-2.07	-.69	1229.77	1.00	*
Relatedness × Emotion (positive vs. negative) > 0	1.44	.63	.40	2.48	79.81	.99	*
Negative condition as the reference level							
Relatedness < 0	.55	.45	-.18	1.29	8.22	.89	

Note: A star indicates if 0 lies outside the 95% CI.

were found in either the negative ($r = .22$, $p = .24$; see Figure 4(C)) or neutral ($r = .27$, $p = .15$; see Figure 4(D)) prosody conditions.

4. Discussion

The present study examined how emotional prosody influences lexicalisation of newly-learned words learned through VR. Emotional prosody was manipulated and learning outcomes were assessed using both behavioural

and ERP measures. The results showed that newly learned words were successfully lexicalised after two days of learning, as indicated by robust behavioural semantic priming effects. Moreover, these effects were associated with learners' subjective emotional experience in the VR environment, especially for words learned with positive prosody. Critically, positive prosody facilitated lexicalisation, as evidenced by larger semantic priming effects in both behavioural performance and the LPC. We elaborate on these findings in the next sections.

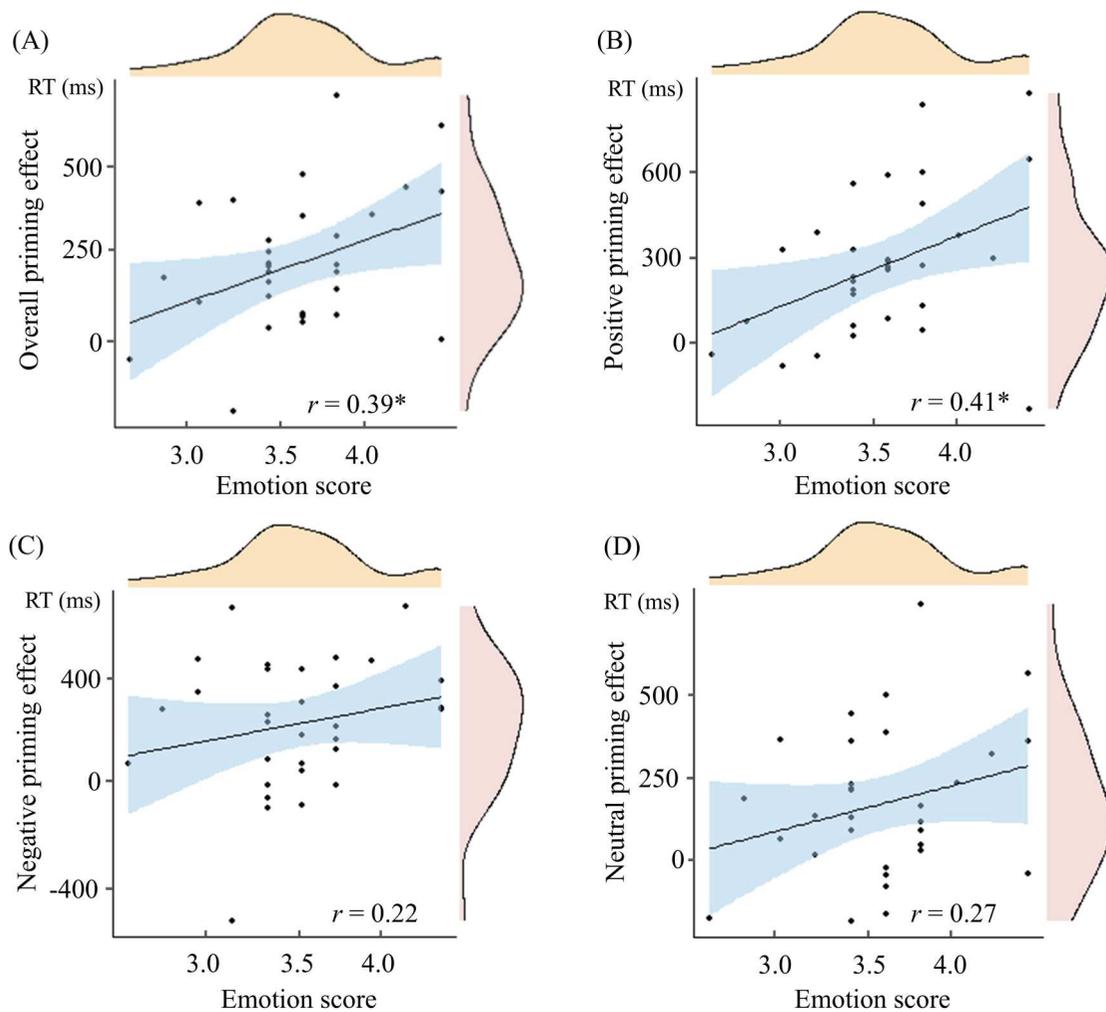


Figure 4. An increasing trend in subjective emotion scores was associated with an overall semantic priming effect (A), and specifically in positive conditions (B), but not in negative (C) or neutral (D) conditions. Blue areas represent 95% confidence intervals.

4.1. Lexicalisation of newly-learned words learned through VR

Words learned through VR showed evidence of lexicalisation, as reflected by participants' significantly faster responses to related prime-target pairs compared to unrelated pairs (i.e. a semantic priming effect) across positive, neutral, and negative conditions. These findings align with previous research on lexicalisation of newly-learned words (e.g. Tamminen & Gaskell, 2013), which has similarly reported semantic priming effects after a 24-hour time period. This suggests that offline lexicalisation may be a central feature of word acquisition, contributing to the integration of new words into existing lexical networks and providing empirical evidence for the Complementary Learning Systems model (Liu & van Hell, 2020). Moreover, the robustness of the observed lexicalisation effects was underscored by a modality shift between learning and testing. Specifically, while the learning phase occurred

in a visual-auditory modality (interacting with 3D objects and hearing German words), the testing phase was a reading-auditory modality (reading Chinese primes and hearing German targets). This design requires cross-modal transfer, ensuring that the priming effects reflected word lexicalisation rather than superficial visual associations (e.g. matching a sound to a specific visual image). These effects indicate that newly-learned words were not bound to the specific VR context, but rather were effectively integrated into learners' mental lexicon.

While all newly-learned words were successfully lexicalised at the behavioural level, only those learned with positive prosody elicited a LPC semantic priming effect, whereas no N400 effects were observed for any condition. This finding is consistent with previous studies on word learning paired with verbal definitions (Lei et al., 2022; Liu & van Hell, 2020), in which only LPC (but not N400) semantic priming effects for newly-learned

words were observed in offline lexicalisation after one day. Assuming that the N400 component indexes automatic aspects of lexicalisation, and the LPC indexes more strategic/explicit and controlled aspects of lexicalisation, this finding suggests that lexicalisation of newly-learned words is a gradual process which shifts from controlled to automatic processing (Bakker et al., 2015; Liu & van Hell, 2020). One or two days of learning is insufficient for the new words' meanings to become automatically lexicalised, which may explain the absence of N400 effects. In sum, the presence of significant behavioural and LPC priming effects, and the absence of N400 effects, indicates that the participants had successfully lexicalised the words to a degree that allowed for controlled semantic access, even if automatic integration had not yet been established.

We also found a significant correlation between learners' self-reported emotion scores and the overall semantic priming effect (averaged across the three types of emotional prosody). This suggests that the learners' intrinsic emotional experience (i.e. experiencing more happiness) in the VR environment predicted the degree of lexicalisation. More interestingly, this intrinsic emotion effect interacted with positive prosody, but not with negative or neutral conditions. Based on the ICALM framework (Plass & Kaplan, 2016) and the affective facilitation hypothesis (Koo et al., 2012), positive emotions arising from both intrinsic and external channels may enhance learners' satisfaction and interest, thereby benefiting cognitive processing and learning. This includes promoting learners' motivation and allocating more cognitive resources to target stimuli. Thus, in a VR learning environment, both positive self-feelings and emotional prosody appear to facilitate lexicalisation of newly-learned words.

4.2. Emotional prosody affects lexicalisation of newly-learned words through VR

Regarding the effect of emotional prosody on lexicalisation of newly-learned words, a critical finding was that there was a larger behavioural semantic priming effect on words learned in conditions of positive prosody. This suggests that positive prosody facilitated the lexicalisation of new words learned in the VR context. There are several theoretical accounts that help explain this finding. Based on the CAMIL (Makransky & Petersen, 2021), we speculate that such facilitation may stem from the fact that positive prosody can improve learners' motivation and interest. In the context of the ICALM (Mayer, 2020), the facilitation highlights the critical role of affective processing in multimodal learning. Moreover, this finding also corroborates the *Positivity*

Principle in multimodal learning, which states that positive emotional elements in learning materials can enhance learners' positive emotional feelings and facilitate their learning (Lawson & Mayer, 2022; Zhao & Mayer, 2025).

To better understand the facilitative effect of positive prosody on lexicalisation of newly-learned words, we examined the breakdown of RTs. We found that enhanced priming in positive conditions was driven by a slowing of unrelated trials rather than an acceleration of related trials. Indeed, while conditions of positive prosody slowed down performance on related trials slightly (+50 ms compared to neutral conditions), it slowed down unrelated trials substantially (+164 ms compared to neutral conditions). In other words, RTs in the positive-related condition were not significantly faster than in the neutral-related condition. These findings suggest that processing emotional prosody (i.e. positive prosody) may capture attention and compete with task-related processing such as semantic judgments (Gupta, 2019; Oliveira et al., 2013), leading to a general slowing of responses. However, despite this general processing cost, positive prosody significantly amplified the magnitude of the semantic priming effect ($RT_{unrelated} - RT_{related}$). This finding indicates that while positive emotion does not facilitate absolute retrieval speed, it enhances the strength of semantic integration (i.e. lexicalisation), resulting in a more pronounced differentiation between related and unrelated trials.

At the neurocognitive level, the ERP findings suggest that lexicalisation of newly-learned words through VR was modulated by emotional prosody. Specifically, words learned with positive prosody elicited a larger LPC semantic priming effect (indexing controlled lexicalisation) compared to words learned with negative prosody, but there was no significant N400 effect (indexing automatic lexicalisation). This finding represents a more elaborated and controlled processing of positive emotions, leading to better lexicalisation of newly-learned words that are learned with positive prosody (Carretié et al., 2008; Hinojosa et al., 2020). As discussed above, lexicalisation of newly-learned words involves lexical-semantic retrieval which gradually shifts from a more controlled to a more automatic process (Bakker et al., 2015). However, this shift is unlikely to occur over two days of learning, hence the lack of significant N400 effects (Lei et al., 2022; Liu & van Hell, 2020).

Another finding was the significant main effect of emotion, showing that responses in positive and negative conditions were slower than neutral conditions and were accompanied by changes in P200 amplitude. In line with previous studies showing enhanced P200

responses on emotional words (Sianipar et al., 2015), the observed changes in P200 amplitude may reflect the additional allocation of attentional resources to analyse emotion in the pre-lexical stage. Moreover, a P200 relatedness effect was only observed for words learned with neutral prosody, but not for those learned with emotional prosody. We speculate that the observed P200 relatedness effect is related to attention, given that the P200 component has been associated with early attentional allocation and prediction (León et al., 2010; Paulmann et al., 2013). Moreover, reading a related prime reduced attentional demands and facilitated prediction during the pre-lexical processing stage, as shown by reduced P200 amplitude for words learned with neutral prosody. By contrast, emotional prosody, both positive and negative, introduced emotional salience that enhanced P200, thereby cancelling out relatedness effects.

4.3. Limitations and future directions

We acknowledge several limitations of the study which, if addressed in future work, may help to establish new research avenues. First, the linguistic stimuli varied in complexity and length, ranging from simple words like *Apfel* to compound words like *Wassermelone*. While the within-item design (comparing the same target across related and unrelated conditions) mitigates the impact of differences in physical stimuli on priming effects, such variation in word length may have introduced latency jitter in the grand average ERPs, particularly for late components. Future studies should aim to include a larger set of strictly controlled, length-matched stimuli to reduce this temporal variability.

Second, the current study did not include a non-VR control group (e.g. a comparison group that learned new words through picture-word association). Consequently, while our findings demonstrate how emotional prosody functions within an immersive VR environment, we cannot causally attribute the observed learning benefits solely to this specific learning context. Future research should directly compare VR and non-VR modalities to isolate the specific contributions of immersive learning to the effects of emotional prosody on lexicalisation.

Finally, regarding the time-course of lexicalisation, we did not observe the classic N400 semantic priming effect, which is typically associated with automatic semantic integration. This suggests that after two days of training, newly-learned words were likely processed via controlled mechanisms (reflected by the LPC) rather than being automatically accessed. It would be beneficial in future work that longitudinal studies track

the consolidation process over a longer period to observe the transition from controlled (LPC) to automatic (N400) processing.

5. Conclusion

The current study investigated the role of emotional prosody in newly learned word lexicalisation in a VR-assisted context. Both behavioural and EPR results support a learning benefit in presenting positive prosody for novel word lexicalisation, and intrinsic emotional experience in VR could modulate the emotional prosody effect on lexicalisation. Our study provides empirical support for the ICALM, demonstrating that emotional prosody influences lexicalisation of words recently learned through VR.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The study was supported by the National Social Science Fund of China (25BY080).

References

- Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, 79, 33–41. <https://doi.org/10.1016/j.neuropsychologia.2015.10.020>
- Blakeman, S., & Mareschal, D. (2020). A complementary learning systems approach to temporal difference learning. *Neural Networks*, 122, 218–230. <https://doi.org/10.1016/j.neunet.2019.10.011>
- Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology*, 55(6), e13049. <https://doi.org/10.1111/psyp.13049>
- Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, 17(4), 628–639. <https://doi.org/10.1037/emo0000263>
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Bürkner, P. C. (2021). Bayesian Item Response Modeling in R with brms and Stan. *Journal of Statistical Software*, 100(5), 1–54. <https://doi.org/10.18637/jss.v100.i05>
- Carretié, L., Hinojosa, J. A., Albert, J., López-Martín, S., de la Gándara, B. S., Igoa, J. M., & Sotillo, M. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology*, 45(2), 188–196. <https://doi.org/10.1111/j.1469-8986.2007.00617.x>

- Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability*, 14(6), 3147. <https://doi.org/10.3390/su14063147>
- Chiu, T. K., & Churchill, D. (2016). Design of learning objects for concept learning: Effects of multimedia learning principles and an instructional approach. *Interactive Learning Environments*, 24(6), 1355–1370. <https://doi.org/10.1080/10494820.2014.948450>
- 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>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological Bulletin*, 128(2), 243–249. <https://doi.org/10.1037/0033-2909.128.2.243>
- Escudero, P., Jones Diaz, C., Hajek, J., Wigglesworth, G., & Smit, E. A. (2020). Probability of heritage language use at a supportive early childhood setting in Australia. *Frontiers in Education*, 5, 93. <https://doi.org/10.3389/educ.2020.00093>
- Fuhrman, O., Eckerling, A., Friedmann, N., Tarrasch, R., & Raz, G. (2021). The moving learner: Object manipulation in virtual reality improves vocabulary learning. *Journal of Computer Assisted Learning*, 37(3), 672–683. <https://doi.org/10.1111/jcal.12567>
- Fyfield, M., Henderson, M., & Phillips, M. (2022). Improving instructional video design: A systematic review. *Australasian Journal of Educational Technology*, 38(3), 155–183. <https://doi.org/10.14742/ajet.7447>
- Gelman, A., & Carpenter, B. (2020). Bayesian analysis of tests with unknown specificity and sensitivity. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 69(5), 1269–1283. <https://doi.org/10.1111/rssc.12405>
- Guo, J., Zou, T., & Peng, D. (2018). Dynamic influence of emotional states on novel word learning. *Frontiers in Psychology*, 9, 537. <https://doi.org/10.3389/fpsyg.2018.00537>
- Gupta, R. (2019). Positive emotions have a unique capacity to capture attention. *Progress in Brain Research*, 247, 23–46. <https://doi.org/10.1016/bs.pbr.2019.02.001>
- Hinojosa, J. A., Moreno, E. M., & Ferré, P. (2020). Affective neurolinguistics: Towards a framework for reconciling language and emotion. *Language, Cognition and Neuroscience*, 35(7), 813–839. <https://doi.org/10.1080/23273798.2019.1620957>
- Jiao, L., Lin, Y., Schwieter, J. W., & Liu, C. (2026). Learning novel words in an immersive virtual-reality context: Tracking lexicalization through behavioral and event-related-potential measures. *Language Learning*, 76(1), 5–27. <https://doi.org/10.1111/lang.12707>
- Jiao, L., Liu, C., Schwieter, J. W., & Chen, B. (2021). Switching between newly learned languages impacts executive control. *Psychophysiology*, 58(10), e13888. <https://doi.org/10.1111/psyp.13888>
- Jiao, L., Zhu, M., Xu, Z., Zhou, G., Schwieter, J. W., & Liu, C. (2024). An ERP study on novel word learning in an immersive virtual reality context. *Bilingualism: Language and Cognition*, 27(1), 128–136. <https://doi.org/10.1017/S136672892300038X>
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2), 163–178. <https://doi.org/10.1111/1469-8986.3720163>
- Jungers, M. K., Hupp, J. M., Rardon, J. A., McDonald, S. A., & Song, Y. (2024). The effect of emotional prosody and referent characteristics on novel noun learning. *Language and Cognition*, 16(4), 1881–1898. <https://doi.org/10.1017/langcog.2024.31>
- Kim, S. K., & Sumner, M. (2017). Beyond lexical meaning: The effect of emotional prosody on spoken word recognition. *Journal of the Acoustical Society of America*, 142(1), EL49–EL55. <https://doi.org/10.1121/1.4991346>
- Koo, M., Clore, G. L., Kim, J., & Choi, I. (2012). Affective facilitation and inhibition of cultural influences on reasoning. *Cognition & Emotion*, 26(4), 680–689. <https://doi.org/10.1080/02699931.2011.595509>
- Lago, S., Stone, K., Oltrogge, E., & Veríssimo, J. (2023). Possessive processing in bilingual comprehension. *Language Learning*, 73(3), 904–941. <https://doi.org/10.1111/lang.12556>
- Lawson, A. P., & Mayer, R. E. (2022). Does the emotional stance of human and virtual instructors in instructional videos affect learning processes and outcomes? *Contemporary Educational Psychology*, 70, 102080. <https://doi.org/10.1016/j.cedpsych.2022.102080>
- Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages*, 4(1), 13. <https://doi.org/10.3390/languages4010013>
- 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>
- León, I., Díaz, J. M., de Vega, M., & Hernández, J. A. (2010). Discourse-based emotional consistency modulates early and middle components of event-related potentials. *Emotion*, 10(6), 863–873. <https://doi.org/10.1037/a0020467>
- 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>
- Liu, C., Feng, J., Schwieter, J. W., Geng, J., & Jiao, L. (2026). Neural correlates of novel word learning in an immersive virtual reality environment. *Journal of Neurolinguistics*, 78, 101300. <https://doi.org/10.1016/j.jneuroling.2025.101300>
- Liu, C., Mao, Y., Wang, X., Schwieter, J. W., & Jiao, L. (2024). Sleep-dependent consolidation effects on foreign language word acquisition in a virtual reality environment. *Memory & Cognition*, 52(2), 302–311. <https://doi.org/10.3758/s13421-023-01461-z>
- 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.12410>
- Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical

- research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33(3), 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Mauchand, M., & Pell, M. D. (2021). Emotivity in the voice: Prosodic, lexical, and cultural appraisal of complaining speech. *Frontiers in Psychology*, 11, 619222. <https://doi.org/10.3389/fpsyg.2020.619222>
- Mayer, R. E. (2020). Searching for the role of emotions in e-learning. *Learning and Instruction*, 70, 101213. <https://doi.org/10.1016/j.learninstruc.2019.05.010>
- Mayer, R. E. (2021). Cognitive theory of multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57–72). Cambridge University Press. <https://doi.org/10.1017/9781108894333.008>
- 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>
- Milne, A. J., & Herff, S. A. (2020). The perceptual relevance of balance, evenness, and entropy in musical rhythms. *Cognition*, 203, 104233. <https://doi.org/10.1016/j.cognition.2020.104233>
- Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>
- Murakami, A., & Ellis, N. C. (2022). Effects of availability, contingency, and formulaicity on the accuracy of English grammatical morphemes in second language writing. *Language Learning*, 72(4), 899–940. <https://doi.org/10.1111/lang.12500>
- Oliveira, L., Mocaiber, I., David, I. A., Erthal, F., Volchan, E., & Pereira, M. G. (2013). Emotion and attention interaction: A trade-off between stimuli relevance, motivation and individual differences. *Frontiers in Human Neuroscience*, 7, 364. <https://doi.org/10.3389/fnhum.2013.00364>
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797. <https://doi.org/10.1037/edu0000241>
- Paulmann, S., Bleichner, M., & Kotz, S. A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, 4, 345. <https://doi.org/10.3389/fpsyg.2013.00345>
- 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/s13428-017-0914-3>
- Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In S. Tettegah & M. Gartmeier (Eds.), *Emotions, technology, design, and learning* (pp. 131–161). Elsevier. <https://doi.org/10.1016/B978-0-12-801856-9.00007-4>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R foundation for statistical computing, Vienna, Austria. <https://www.R-project.org/>
- Schirmer, A., & Kotz, S. A. (2003). Erp evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience*, 15(8), 1135–1148. <https://doi.org/10.1162/089892903322598102>
- Sianipar, A., Middelburg, R., & Dijkstra, T. (2015). When feelings arise with meanings: How emotion and meaning of a native language affect second language processing in adult learners. *PLoS One*, 10(12), e0144576. <https://doi.org/10.1371/journal.pone.0144576>
- Sumner, M., Kim, S. K., King, E., & McGowan, K. B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology*, 4, 1015. <https://doi.org/10.3389/fpsyg.2013.01015>
- 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/10.1080/17470218.2012.724694>
- Van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. *The European Health Psychologist*, 16(2), 75–84. <https://www.ehps.net/ehp/index.php/contents/article/view/ehp.v16.i2.p75>
- Zhang, J., Wu, C., Yuan, Z., & Meng, Y. (2020). Different early and late processing of emotion-label words and emotion-laden words in a second language: An ERP study. *Second Language Research*, 36(3), 399–412. <https://doi.org/10.1177/0267658318804850>
- Zhao, F., & Mayer, R. E. (2025). Improving multimedia learning with emotional design: Depicting key elements with positive features. *Journal of Computer Assisted Learning*, 41(3), e70028. <https://doi.org/10.1111/jcal.70028>