



A computational model of TE-dominant noticing, repetition, prior knowledge and grammatical knowledge acquisition

Juan Yang^{1,4} · X. F. Qi² · R. Liu¹ · L. Wang¹ · B. Sun³

Accepted: 22 February 2022
© The Author(s) 2022

Abstract

Computer-assisted textual enhancement (CATE) technology has been widely used to improve English as foreign language (EFL) learners' syntactical and grammatical learning. Visual attention, repetition, and prior knowledge are known as the vital factors in CATE-assisted knowledge-acquisition; however, there still lacks a model which can describe those factors' intrinsic cooperating-mechanism that works in the CATE-based knowledge-acquisition. Therefore, this paper built up a computational model (PESE) of using those factors as variables, by fitting and predicting the data collected from empirical experiments with an average accuracy of 78%, PESE testified and complemented the assumptions proposed by previous studies. PESE suggested that although the efficacy of CATE is majorly decided by learners' prior-knowledge of the targets, the interactive effects of visual-attention, repetition, and inductive activity could partly compensate for the effect from prior-knowledge, and the efficacy ceiling of repetition also could be estimated according to the 'easy-perceiving level' coefficient. At the end of this paper, 3 pedagogical implications were proposed for English teachers who are willing to integrate CATE into their teaching activities.

Keywords Computer-assisted textual enhancement (CATE) · English as foreign language (EFL) · Computational model · Visual attention · Prior knowledge · Repetition

X. F. Qi joint first author.

✉ Juan Yang
Jkxy_yjuan@sicnu.edu.cn

¹ College of Computer Science, Sichuan Normal University, Chendu, China

² School of Education, Durham University, Durham, UK

³ College of Foreign Languages, Sichuan Normal University, Chendu, China

⁴ Department of Educational Technology, Sichuan Normal University, Chendu, China

Introduction

Improved syntactic- and grammatical awareness can be beneficial to students in enhancing their reading and writing skills in second language (L2) acquisition. However, it is very hard for young L2 learners to develop a strong knowledge of the syntax to the extent that they can perform syntactic parsing as quickly as they do in their first language (L1). The major obstacle preventing L2 readers from developing responses regarding L2 syntactic features may be the restricted exposure to L2 prints (Park & Warschauer, 2016). Young L2 students, lacking L2 linguistic knowledge and insufficiently exposed to L2 prints, have difficulties in noticing and acquiring the formal features of written discourse, resulting in significantly low academic performances in reading and writing.

One of the approaches commonly used to help students pay sufficient attention to the syntactic information is textual enhancement (Park & Warschauer, 2016) based on the 'Noticing Hypothesis' (Schmidt, 1990). The input features entail typographical cues so that the target form is enhanced and its visual appearance in the text is altered (italicised/boldened/underlined, see Kim, 2006). The aim of the typographical change is to enhance the target form's perceptual salience, e.g., using colours to improve the perception of the linguistic cues (Comeaux & McDonald, 2018), using automatic colour font to enhance the perception of the articles (Ziegler et al., 2017), using colour, bold, and underlining to help acquire verbal morphology (Mayén, 2013), using underlining to enhance the perception of collocations (Szudarski & Carter, 2016) and the memorisation of target words (Boers et al., 2017), using visual-syntactic text formatting (VSTF) technology to attract learners' attention not only to a single feature, but also to a syntactic structure (Park & Warschauer, 2016). The syntactic features omitted by naturalistic L2-learners could be prominent by reforming learners' attention distribution (Cintrón-Valentín & Ellis, 2016). Up to now, the effects of textual enhancement on grammatical learning or syntactic-knowledge acquisition have been widely acknowledged (e.g., Gascoigne, 2006; Lee & Huang, 2008; Smith, 1993). Moreover, with the fast development of information technology, computer-assisted textual enhancement (CATE) has become the dominant form in improving learners' apperception of the syntactical and grammatical learning (e.g., Comeaux & McDonald, 2018; LaBrozzi, 2016; Park et al., 2012; Park & Warschauer, 2016; Szudarski & Carter, 2016; Winke, 2013).

Visual attention was reported as the corner stone of applying textual enhancement technologies, given that it is the prerequisite of transferring L2 input to intake (e.g., see Comeaux & McDonald, 2018 for review). In addition to visual attention, many other factors were also reported to be of vital importance in deciding the efficacy of applying CATE to L2 knowledge learning, e.g., Lee and Huang (2008) claimed that the CATE efficacy was a function of the learners' prior knowledge, and Szudarski and Carter (2016) implied that the overloaded repetition would lead learners into a resistance period of perceiving the CATE-based intervention.

Previous studies had researched those factors' dependent impacts on the efficacy of CATE, finding that the effectiveness of CATE was influenced by the

interaction of those factors, e.g., Comeaux and McDonald (2018) suggested that the interaction of the type of the cues, the cues' appearing frequency, and the dominance of the cues affected the learning effects of CATE-based intervention. Actually, as early as in 2008, Han et al. (2008) had systematically discussed the possible interactions among CATE, prior knowledge and repetition based on empirical studies (for review, see Han et al., 2008). Their analysis laid the theoretical foundation of revealing the intrinsic cooperating mechanism of those factors that work within a CATE-based intervention.

Literature review

CATE in L2 grammatical and syntactic knowledge acquisition

Grammatical- and syntactic awareness, is described as the ability of reflecting on sentence structures and the related rules of using and arranging words (Park & Warschauer, 2016). In the last 10 years, CATE technique has been widely developed and used in increasing the flexibility of the method, because the multi-media technology could make the computer-mediated contexts more vividly by incorporating sound, animation and interactivity (Gascoigne, 2006; Liu & Leveridge, 2017; Meurers et al., 2010; Park & Warschauer, 2016). Most importantly, multi-media enables learners to generalise abilities of exploiting the textual, aural, and visual cues in developing comprehension (Ulitsky, 2000).

Most successful cases of applying CATE appeared in grammatical and syntactic knowledge learning. The effects brought by CATE in grammar learning were widely acknowledged. For example, Park et al. (2012) suggested in their study that CATE could effectively increase the attention allocation on targets. Consequently, the participants in the team with CATE assistance outperformed their peers who were without CATE assistance. In addition to the single syntactic-feature acquisition, CATE was also used to improve the noticing of the syntactic structures (Park & Warschauer, 2016). In Walker et al.'s study (2005), they employed VSTF technology to visually separate and indent the phrases that were nested within a single sentence, and brought the meaning and underlying structures of a given sentence to a more prominent position. Their experimental results showed that low-proficiency students gained significant improvements in syntactic awareness. Nevertheless, it was also reported that CATE may impede reading comprehension, e.g., in Park et al.'s study (2012), they implied that the attention allocated to enhanced targets reduced the processing of comprehension. Winke (2013) similarly concluded the trade-off of using CATE in second language acquisition (SLA), that is "More attention to forms resulted in less to meaning."

In sum, the successful cases mentioned above shed light on how to assist learners in noticing and comprehending syntactic knowledge that is hidden behind combinations of words, that is, the syntactic features omitted by naturalistic L2-learners could be prominent by reforming learners' attention distribution (Cintrón-Valentín & Ellis, 2016).

Visual attention, prior knowledge, and repetition in CATE

Exposing adult SLA beginners to intensive input to generate stable recognition patterns regarding L2 representation is advocated by Rast (2008), especially to generate awareness of noticing the syntactic features that are hard to be noticed by learners, such as morphological cues with poor semantic meanings (Ellis & Sagarra, 2011). Ellis suggested that, the remedy of the “fossilized” responses toward L2 input is to bring the information into the light of consciousness; therefore, interventions of consciousness raising or form-focus can help learners notice the cue in the first place (Ellis, 2002). The theory that supports this claim originated from Schmidt’s Noticing Hypothesis (Schmidt, 1990), which holds that input cannot be transformed to intake for processing unless it is noticed, that is, consciously registered.

Attention could also be taught and learned (Indrarathne & Kormos, 2017). Ellis et al. (2011) testified the effectiveness of applying intensive pattern training for learners who studied abroad to help them notice target L2 syntactic features by reshaping their learned attention distribution in L1. The results were promising. However, they also admitted that not all learners have the opportunity to study abroad, suggesting that sufficient pedagogical interventions should be added into routine learning, such as providing feedback on errors and applying CATE technology (Ellis & Sagarra, 2010). Although the view that CATE could effectively attract learners’ attention to the targets and improve the apperception of the syntactic knowledge has been widely acknowledged, e.g., see (Winke, 2013), its efficiency of how it transfers learned attention to feature acquisition remains unclear.

Prior knowledge of the targets may significantly affect the degree of acquisition of the enhanced forms (Lee & Huang, 2008; Webb et al., 2013). Prior knowledge is referred to by many studies as “what learners had already known about the target knowledge,” e.g., see (Lee, 2007) and (Winke, 2013). It was assumed that the learners with intermediate knowledge regarding the target forms would receive the best intervention result, given that the beginners had insufficient prior-knowledge for understanding the new knowledge while the advanced learners had abundant previous exposure to the targets (Park et al., 2012). The learners’ prior knowledge was also presumed to be correlated with their understanding of the CATE-targeted knowledge, e.g. Szudarski and Carter (2016) reported that their participants who had rather low vocabulary knowledge were lack of the gains in the form-recall test, although they had noticed the targets; and Park and Warschauer (2016) reported that the enhanced grammatical structures of a sentence were too complicated to be understood for young L2 beginners because of lacking the necessary prior knowledge.

Repetition refers to the enhancing of the salience of target knowledge with a manually increased frequency (Han et al., 2008). Similar to the training epoch in machine learning, repetition has been verified as an important factor for the effectiveness of input flood (Ellis, 2002; Rott, 2007; Yang et al., 2019). However, the effect of repetition is also reported to be limited, e.g., Szudarski and Carter (2016) suggested in their study that the overloaded repetition would not improve learners’ performance, but led them into resistance of perceiving from the intervention; and Rott (2007) found that four times repetition of the textual enhanced content showed no priorities over the CATE content without repetition, but the repetition followed

by CATE showed a prominent efficacy. Moreover, some explicit instructions could override the effects of repetition (Eckerth & Tavakoli, 2012), such as the effect of ‘Motivational-Cognitive Involvement’ (Laufer & Hulstijn, 2001). Although numerous studies, e.g., Han et al. (2008), had already discussed the possible interaction among prior knowledge, repetition, and attention in diverging the effect of CATE-based intervention, questions like, how the effect of repetition varies because of the changing of the attention and prior knowledge, was rarely discussed in the existing studies.

Visual-attention related computational models in L2 learning

Modelling humans’ visual attention in an interactive environment has a long research history. In those computational models, humans are usually actively engaged in a task that is highly correlated with the using of visual-attention, and thus are strongly top-down driven (Borji et al., 2011; Borji & Itti, 2013). Computational models involving visual-attention in L2 learning not only revealed the mechanisms of how people dominated their visual attention but also strove to understand the causal mechanisms that could explain the relationship between a series of observed phenomena (Mareschal & Thomas, 2007). For example, Monaghan et al. (2017) employed a function of L2 language exposure to simulate the weakening of the frequency effect; Yang et al. (2022) built a computational model to simulate the effective boundary of audio-assisted reading with cognitive-load as the modulating variable. This type of computational models focuses on building an explanatory framework for language learning processes as well as interactive factors (for review, see Dijkstra et al., 2019), which even unifies bilingualism within a single infrastructure if the transparency of the model is sufficient (e.g., see Monaghan et al., 2017).

Another usage of building computational models in language learning is to work out unexpected implications of a theory, because “the world is highly interactive, even a simple process theory can lead to unforeseen behaviors.” (Mareschal & Thomas, 2007, p. 2). For this usage, transparency is no longer a prerequisite of the model, as simulating and predicting the unexpected results based on the experience data become major concerns. For example, Wang et al., employed convolutional neural network (CNN; 2018) and Recurrent Neural Network (RNN; 2019) to predict the eye movements of readers reading a previously unseen text by using existing eye movements as training data. Their models implied that successful readers’ visual attention pattern may not be suitable for the less-skilled readers, which stood on the opposite side of a popular pedagogical method that trains less-skilled readers with their successful peers’ attention pattern to help the former achieve better reading outcomes. Although those models cannot explain how the outcome was generated because of the intervention, they provide researchers with unexpected implications based on experience data.

By summarising the references listed above, it can be concluded that computational models involving visual attention in L2 learning are classified into two types according to their usage: white-box-technique-based models and black-box-technique-based models. For the latter, deep learning models for natural language

processing (NLP), e.g., CNN (Aloysius & Geetha, 2017) and Bidirectional Encoder Representations from Transformers (BERT, Devlin et al., 2019) were the first choices of the researchers as those techniques' predicting accuracy was far superior to that of white-box techniques. However, those models are inappropriate for observing causal-result relations among intervention and outcome unless rules can be extracted from the black-boxes. Conclusively, because of the limited explanatory ability of black-box techniques, most visual-attention related computational models adopted white-box-technique as the basic infrastructure.

This study

As mentioned above, although visual attention, repetition of the CATE-based intervention, and the learners' prior knowledge of the targets are of vital importance in deciding the efficacy of applying CATE in L2-knowledge-acquisition, those factors' intrinsic cooperating mechanism, and their compound impacts on L2-knowledge learning remain unclear. Therefore, this study uses different content modes to represent environments in which CATE-based interventions took place, and uses different levels of prior knowledge, visual attention and repetition as variables.

By building a computational model, this paper aims to answer the following questions:

1. How do prior knowledge, visual attention, and repetition intrinsically cooperate with each other in CATE-assisted L2-knowledge acquisition?
2. How does the efficacy of CATE-based intervention vary due to the change of prior knowledge, visual attention, and repetition?

Method

Participants

This study used a design of three comparison groups to represent three types of content modes in which CATE took place for young Chinese-English students; 38 Mandarin-speaking students aged 8 (mean = 8.03, $sd = 0.56$) participated in the experiment. All participants were in second grade at a primary school in Sichuan province of China. Small gifts were promised to each student as rewards for his/her participation.

Thirteen of them were randomly assigned to group-1 (G1), in which they were intervened with the materials presented in one-by-one mode; fourteen of them were randomly assigned to group-2 (G2), in which they were intervened with the materials presented in batch mode (simulating the flood mode); the rest ($n = 11$) were assigned to group-3 (G3), in which they were intervened with the materials presented in a batch mode and at the same time attached with L1 lexical information.

Research context

Experimental design and data processing flow

This study was designed in a quasi-experimental way. Firstly, a $3 \times 3 \times 3 \times 3$ MANOVA between subjects' design was employed by the experiment to observe the phenomena from a behavioural perspective, in which prior-knowledge (poor; general; good), salience (poor; general; good), effort (poor; normal; good), and epoch (1; 2; 3) were the variables. Here, data of prior-knowledge were collected from the pre-test implemented through a card-combination game; and data of effort were calculated through the participants' dwell time proportion of eye-gaze. Three levels of salience correspond to three experimental groups. Three levels of epoch refer to the different intervention rounds.

Secondly, in addition to analysing dwell time proportion, Kullback–Leibler (KL) divergence was used to compare the degree of similarity among the attention patterns of three experimental groups. Finally, based on the findings revealed by those statistical analysis, a linear model was built to simulate the interactive effects on learning outcome made by prior-knowledge, salience, effort, and epoch.

Targeted grammatical knowledge and involved materials

Most EFL teachers adopt usage-based learning strategies (Elgort, 2017; Lin, 2014) in their regular face-to-face classes as most L1 teachers do. Possessive pronouns are not specifically taught in students' formal English learning classes, but are repeatedly confronted during the regular reading, belonging to the grammar that form the foundation of the students' future English learning. However, because of the lack of explicit instructions from teachers, students in grade-2 cannot generate an appropriate correspondence between the personal pronouns and their possessive pronouns in L2. Although participants in [this study](#) had little knowledge of the target forms, they had been repeatedly exposed to the declarative sentences and had the ability to translate subject, verb, and object (SVO) structured sentences into Chinese, i.e., they had acquired SVO declarative sentences on recognition level. Thus, this study chose the transformation from personal to possessive pronouns as the grammatical objective of the tasks to drive participants to deploy their visual attention, using the CATE technique (animation and colour) to make the target grammatical knowledge salient.

Therefore, the materials used in the experiment included two major types of sentence structures: declarative sentences using personal pronouns as subjects and the sentences with possessive pronouns followed by nouns as subjects. The syntactic structures of the sentences were basic ones following the subject, verb and object (SVO) structure. The words were those that had been repeatedly confronted by the participants, and the sentences were listed in "[Appendix A](#)".

Computer-Assisted Textual Enhancement (CATE) design

Approaches commonly used to help students raise attention on targets include discrete-item exercises to increase metalinguistic knowledge, simplifying texts to reduce syntactic complexity, and CATE to make language input salient (Park & Warschauer, 2016). Therefore, after controlling for the difficulty level of the materials and the possible influences from teachers, CATE technology was designed to guide participants' attention towards the target grammatical knowledge.

To reveal the nature of attention that may influence the acquiring of L2 grammatical knowledge from reading, this study designed three types of content presentation modes for simulating three typical reading scenes (see Fig. 1) in which the salience of the CATE was different in each situation. The first type of content mode presents sentences one by one, the enhancement of the target knowledge in each sentence is implemented through animation-based CATE. This content mode was designed to simulate the reading situation, in which students focus on a single sentence each time, and the salience of the CATE is most prominent compared to the rest two content modes. Thus, this content mode will be referred to later as the “good” level of the variable salience, and will be applied in G1. The second type of content mode is called batch-content-presentation mode, and is designed to simulate another reading scene, in which participants may intensely confront similar sentences within a reading content. In this reading situation, the salience of the CATE degrades compared to the former mode, but was the most commonly confronted situation for L2 reading. Thus, this content mode will be referred to as the “general” level of the variable salience, and will be applied in G2. The third type of the content mode was designed to simulate the reading scene that readers label L1 lexical information on the reading contents, which was crowded with L1 information. This reading mode makes the

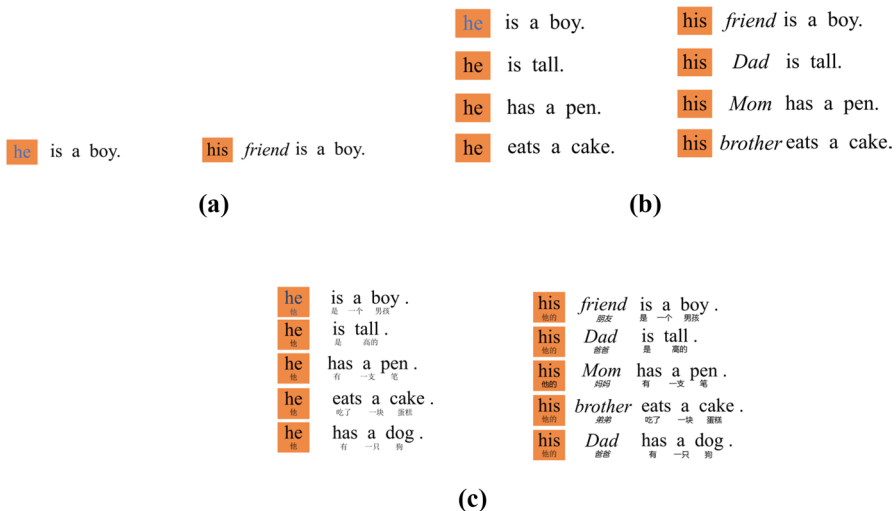


Fig. 1 a Content presentation mode for G1; b Content presentation mode for G2; c Content presentation mode for G3

salience of the CATE the least prominent compared to the other two content modes. Thus, this mode will be referred to as the “poor” level of the variable salience, and will be applied in G3.

Card-combination game

A card-combination game was used in the pre-test, as well as in all post-tests immediately after the intervention with the aim of evaluating participants’ current knowledge of understanding the possessive formation of subject nouns. There were two reasons for choosing the card-combination game in [this study](#): (1) it is a transformation of the sentence production process (Alanen, 1995) and was more like a game than a test for eight-year-old children, and (2) it can collect multi-faceted information to establish a knowledge database for each participant via recording their performance.

In the card-combination game, each participant was asked to generate two independent declarative sentences by selecting out appropriate words given by a teacher. Figure 2 demonstrates the process of forming two sentences by a student in the card-combination game. The first sentence was a declarative one with a personal noun as subject, and the second sentence required the student to use the possessive pronoun as a subject. Teachers gave scrabble cards to a participant and then asked him/her to use all given cards to separately form two complete sentences. All participants’ performances were video recorded. The sentence pairs used in the instant post-tests were¹: (1) he has an apple. his Dad has an apple. (2) She eats a cake. Her Mom eats a cake. (3) It is a ball. Its colour is red.

If the learners generated correct answers for each sentence pair, they scored a point. Otherwise, they scored zero. Therefore, the students’ score range would be [0, 3]. Here, the correct answer was defined as ‘the participants can correctly use possessive pronouns followed by nouns as subjects in the sentences.’ Therefore, student k was tested by three sentence pairs (it-its, she-her, and he-his), and the normalised performance in pre- and each post-test was logged as proportion correct parameter a_k , where $a_k = \frac{\text{score}}{3}$, $a_k \in [0, 1]$.

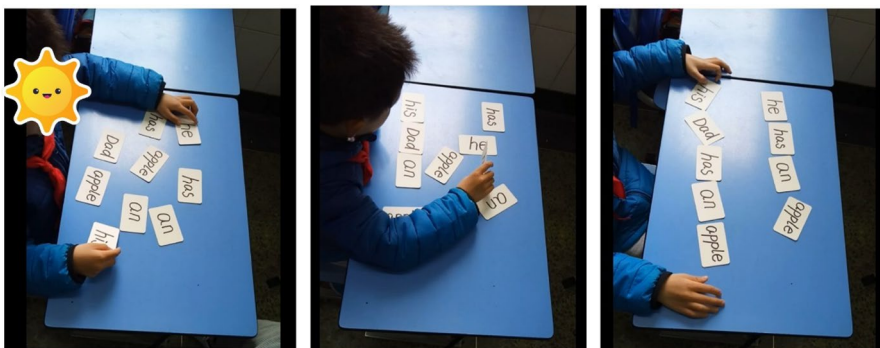


Fig. 2 Screenshots of the beginning, middle, and end of a card-combination game

Measures used to quantify attention

Eye-movement tracking

An eye-tracker (a Tobii T120 running python packages) was employed to monitor students' attention distribution while they received interventions. In our experiment, participants were required to watch videos displayed on the screen of an eye-tracker. The dividing of the different areas of interest is a partition of the screen, as shown in Fig. 3. That is, the union of all areas form a visual domain, and the intersection of the three areas is empty. Three areas are: the area covers target syntactical knowledge (targeted area), the area covers the rest of the content (un-targeted area), and the area that covers the other areas (other area, i.e., blank area and the area outside the screen). The three areas form a partition of the whole visual domain, each one is independent of another. Thus, the attention distribution in each test was formed by participants' eye-gaze dwell proportion on each area.

The effort (E1) was calculated by adding the attention distribution on the target area (Ef1) and on the un-targeted area (Ef2). Original SEEV models defined effort as the physical movement of the human which may inhibit scanning. Thus, effort in the SEEV models may negatively influence the allocation of attention. In many other studies (e.g., Gollan & Ferscha, 2016), effort was interpreted as the participants' effort to attend or engage in a particular task. In these works, effort positively influenced attention allocation. In this paper, students had no physical interactions with the learning environment and thus, the effort can be interpreted as their attention-control ability (attention distribution on the other area). We transferred the effort from negative to positive by calculating the attention distribution spent on reading contents (Ef1 + Ef2), as the summation of the three (Ef1 + Ef2 + eye-gaze dwell proportion on other area) was equal to 1.

The eye-tracking data was calculated to form attention distribution by summing up all the Dwell Time (> 300 ms) regarding a same area of interest in an intervention. The Summed Dwell Time on each area of interest was normalised

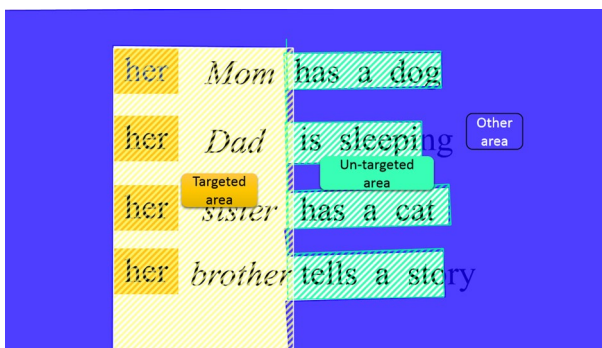


Fig. 3 The dividing of the areas of interesting

into [0, 1] range by computing its proportion to the total Summed Dwell Time covering all areas of interest in each intervention round.

Method of comparing attention distributions

Kullback–Leibler (KL) divergence, also called relative entropy, was employed in this study to compute the mutual similarity among the visual attention distributions of three experimental groups. KL divergence is a commonly used method to measure visual attention. However, because of its asymmetrical characteristics, many studies in vision research expanded it to symmetrical KL divergence (Borji & Itti, 2013). Both asymmetrical and symmetrical KL were positive. The larger the KL value is, the more different those two distributions are. Therefore, the KL value between two identical distributions would be 0. Specifically, symmetrical KL divergence is computed according to the following formula (Borji & Itti, 2013):

$$D_{KL}(H, R) = \frac{1}{2} \times \sum_k^q \left(H_k \log \left(\frac{H_k}{R_k} \right) + R_k \log \left(\frac{R_k}{H_k} \right) \right) \quad (1)$$

where q refers to the probability space of the attention, in our experiment, the visual area was divided into three parts: targeted, un-targeted and other. Thus, $q=3$. H and R are two different distributions, therefore, $\sum_k^q H_k = 1$ and $\sum_k^q R_k = 1$.

Procedure

The experiment was conducted in participants' second semester at grade-2, and lasted about 3 months (the experiment was performed 3–4 days a week, each time completed within an hour). First, all participants were required to implement the pre-test, and then were randomly clustered into three different groups by excluding the participants who achieved full scores in the tests; Second, the participants from three groups were presented with reading contents in different modes, and their eye-movement data was recorded by the eye-tracker; Third, each participant was required to undergo the instant post-tests. The process beginning from the second step to the third step would be conducted for three epochs. All participants received the intervention and instant post-test in 10 min at a frequency of 1 time per week, which is also the acceptable intervention frequency by English teachers. The procedure of the experiment was illustrated in Fig. 4.

Results

Learning effects of the target knowledge

The learning effects (proportion correct a) were calculated from the card-combination games. Table 1 showed the results of a $3 \times 3 \times 3 \times 3$ MANOVA between subjects' design with the variables of prior-knowledge (poor; general; good), salience (poor; general; good), effort (poor; normal; good), and epoch (1; 2; 3).

Fig. 4 The procedure of the experiment

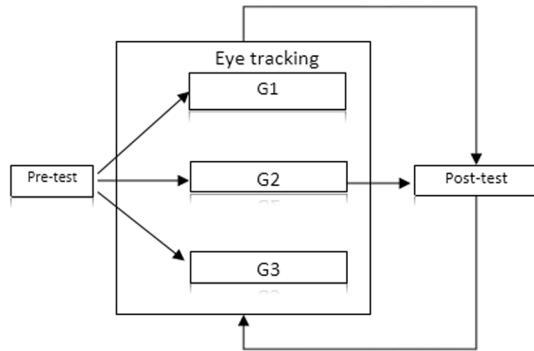


Table 1 3×3×3×3 MANOVA analysis

Source	MSE	F	df	Proportion correct α	
				Sig	η^2
Pre-test (P)	0.978	14.513	2	0.000**	0.372
Saliency (S)	0.063	0.938	2	0.398	0.037
Effort(E1)	0.038	0.569	2	0.570	0.023
Epoch(E2)	0.119	1.765	2	0.182	0.067
S×E1	0.048	0.708	3	0.552	0.042
S×E2	0.025	0.342	3	0.849	0.018
E1×E2	0.037	0.544	4	0.704	0.043
S×E1×E2	0.235	3.486	1	0.068	0.066

**Sig. < .01, $\eta^2 = .01$ (small effect); $\eta^2 = .06$ (medium effect); $\eta^2 = .14$ (large effect); Effort(E1) = attention distribution on targeted area(Ef1) + attention distribution on un-targeted area(Ef2)

Prior knowledge was calculated according to the participant’ pre-test scores, and if $a_k \in [0, 0.33]$, the prior knowledge level of student k was evaluated as “poor”, if $a_k \in (0.33, 0.67]$, the prior-knowledge level of student k was evaluated as “general”, while a_k was located in the area of $(0.67, 1]$, the prior-knowledge level of student k was evaluated as “good.”

Additionally, the effort scores were divided into three areas in a non-increasing order, and defined effort = “good” if the effort score was located in the area of $(1, 0.98]$ area, effort = “normal” if the scores were located in the area of $(0.98, 0.95]$, and effort = “poor” if the scores were located in the area of $(0.95, 0.73]$, where the threshold value 0.95 was the average of all effort values, and the threshold 0.98 was the average of all effort values that are higher than 0.95.

Figure 5 plotted the proportion correct α at pre-test and post-tests for three groups, suggesting that only the improvement for G3 (mean = 0.36) was higher than G1 (mean = 0.25) and G2 (mean = 0.14).

It can be seen from Table 1 that no variable or interactive effects among multi-variables could significantly influence participants’ final proportion correct except for the pre-test score (P), implying that the three commonly confronted

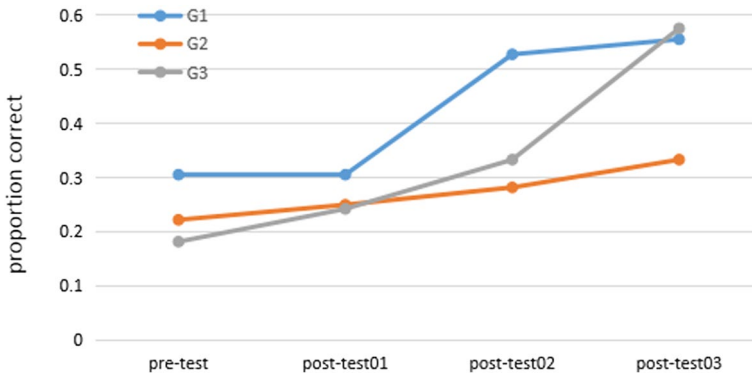


Fig. 5 Proportion correct α at pre-test and 3 post-test points for three groups

reading situations that CATE took place had no significant difference in influencing the acquisition of the current grammatical knowledge. Effort generated from CATE and the intervention epochs had limited effect size in the current task. Although prior knowledge of three groups had no significant difference [LSD(G1,G2)=0.298; LSD(G1,G3)=0.22; LSD(G2,G3)=0.812], the prior knowledge influenced the acquisition results with a large effect size.

Attention distribution patterns

This section compares the mutual KL divergences among three groups and the results are listed in Table 2. Table 2 revealed that the attention distribution of G3

Table 2 Attention distribution and mutual KL divergence among three experimental groups

Intervention epoch	Ef ₁ (targeted area)		Ef ₂ (un-targeted area)		Dwell time proportion of other area		Relative entropy $D_{KL}(G1, G2)$
	G1	G2	G1	G2	G1	G2	
Epoch 1	0.5241	0.5231	0.4481	0.4355	0.0276	0.0412	0.0041
Epoch 2	0.4884	0.5204	0.4536	0.4307	0.0578	0.0480	0.0034
Epoch 3	0.4830	0.5353	0.4504	0.3999	0.0665	0.0646	0.0082
	G1	G3	G1	G3	G1	G3	$D_{KL}(G1, G3)$
Epoch 1	0.5241	0.5035	0.4481	0.4896	0.0276	0.0068	0.0242
Epoch 2	0.4884	0.5273	0.4536	0.4592	0.0578	0.0134	0.0489
Epoch 3	0.4830	0.5083	0.4504	0.4704	0.0665	0.0212	0.0388
	G2	G3	G2	G3	G2	G3	$D_{KL}(G2, G23)$
Epoch 1	0.5231	0.5035	0.4355	0.4896	0.0412	0.0068	0.0497
Epoch 2	0.5204	0.5273	0.4307	0.4592	0.0488	0.0134	0.0342
Epoch 3	0.5353	0.5083	0.3999	0.4704	0.0646	0.0212	0.0441

is significantly different from those of G1 and G2 since $D_{KL}(G1, G2)$ is almost 10 times of $D_{KL}(G1, G3)$ and $D_{KL}(G2, G3)$. L1 information almost reshaped the attention distribution via increasing the effort on searching information.

Table 3 showed the ANOVA analysis for the effort-related variables. It revealed that the content mode significantly influenced learners' physical effort of exploring information that may be related to the task. The influence on the effort spent on target area was insignificant, $F(2, 99)=2.4, p=0.093$, but the influence on the effort spent on the un-targeted area was significant, $F(2,99)=5.04, p=0.008$. Especially when the contents were integrated with L1 information, the effort-devoting was significantly higher than other two attention patterns.

Modelling the prior knowledge, visual attention, and repetition in grammatical knowledge acquisition

So far, we have verified the already known fact for most EFL teachers, that is, the prior knowledge effects as a major factor that influences the effectiveness of using CATE. But, how CATE-dominant effort, the salience of CATE and the repetition tune the learning effects remains unknown. Here, we model the prior knowledge (P), the CATE-dominant effort (E, Ef_1 and Ef_2), the salience of CATE (S), and repetition (E) as Formula (2) shows:

$$P_t(\text{Acq}) = \sum_{k=1}^3 [P_t + b * (Ef_1)(Ef_2)(S_t)(E_2)] \tag{2}$$

where t refers to a particular grammatical knowledge targeted in this study and P_t =prior grammatical awareness, which was updated according to the evolving of the epoch, i.e., the results of the last instant post-test in the last round were considered as the prior grammatical awareness of this round; Ef_1 =attention distribution on targeted area; Ef_2 =attention distribution on un-targeted area; S =Salience; E_2 =Epoch; b was the 'easy-perceiving level' coefficient, which was designed as an estimator of the target knowledge based on the novel assumption that the easier the target knowledge is, the faster the target knowledge be perceived by learners. In this formula, Prior knowledge (P) is considered as the dominant factor, and the interactive influences among Effort, Salience and Epoch (ESE) are considered as one

Table 3 ANOVA analysis for the effort-related variables

	G1		G2		G3		Sig	η^2
	M	SD	M	SD	M	SD		
Ef1	0.486362	0.077310	0.528572	0.082560	0.511600	0.080537	0.093	0.046856
Ef2	0.438099	0.075323	0.411267	0.093751	0.474424	0.074574	0.008**	0.092433
E1	0.924461	0.061307	0.939839	0.038739	0.986024	0.015670	0.000**	0.264435

**Sig. < .01, $\eta^2 = .01$ (small effect); $\eta^2 = .06$ (medium effect); $\eta^2 = .14$ (large effect); Effort(E1)=attention distribution on targeted area(Ef1) + attention distribution on un-targeted area(Ef2)

modulating parameter in the fitting. We employed logistic model which was also adopted by SEEV (Horrey et al., 2006) as the regressor and compressor to simulate the ESE effect. Although the interactive influence among variables were also taken into account to build the model, Level-2, degree-2 polynomial logistic regression model achieved best fitting results as the MSE (mean squared error) of predicting positive samples in both training dataset and cross validation dataset in this model was minimal. It needs to be noted that the nature of the PESE model is a linear model, which follows the principle of transparency of building computational models in explaining how human learns. 80% of the samples from this study’s experiment were used as training data, and 20% of them were considered as cross-validation data.

In Fig. 6a, we plotted the variability in Ef_1 by Ef_2 of three content-presentation modes in this study, as well as the variability in Ef_1 by Ef_2 in acquiring the ‘be’-question features. This figure represented a mutual-relationship between Ef_1 and Ef_2 , showing that the attention spent on targeted content and un-targeted content was mutually against to each other, but the variance of the Ef_1 by Ef_2 was largely due to the salience of the targets.

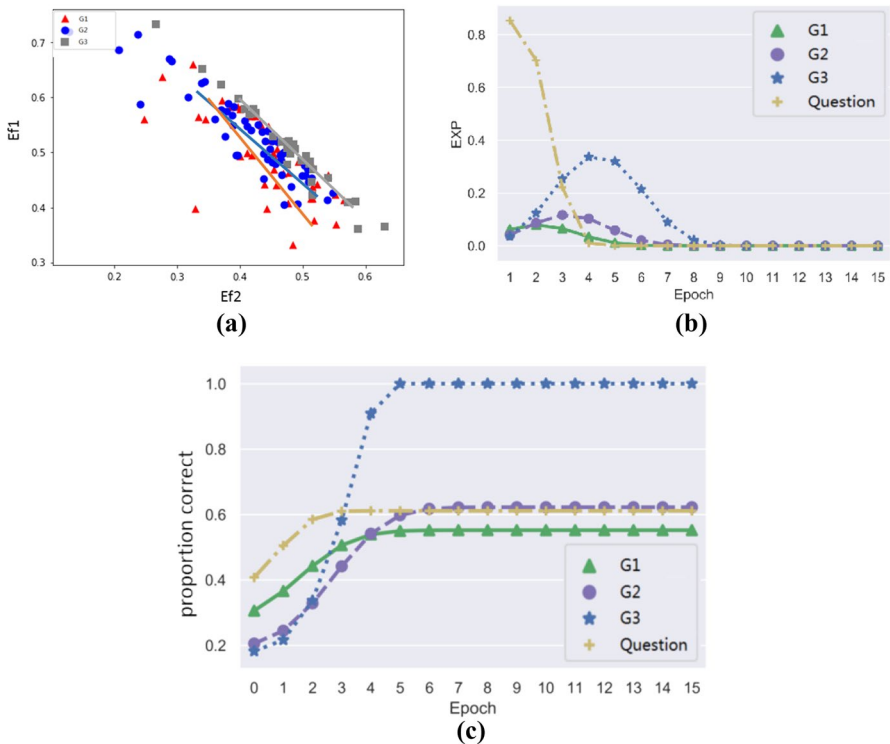


Fig. 6 a Variability in Ef_1 by Ef_2 . b Intervention efficacy of $Exp(Ef_1, Ef_2, S_r, E_2)$ by intervention epochs. c Proportion correct by intervention epochs

To evaluate the robustness of this PESE computational model, we employed another dataset to train and test the model, the data was originated from an experiment of acquiring English questions' features with the assistance of CATE. 80% of the samples from that experiment were used as training data, and the rest 20% of them were used as cross-validation data. The experiment involved a same group of participants and was conducted in a semester before.² The linguistic features involved in that experiment were those of English 'be'-questions, and the experiment also employed animation to enhance the salience of the targets, the content was presented in a batch mode without L1 information. The computing of the accuracy proportion a_k was in line with the normalisation method used in this paper. After trained the model by using two different datasets from two experiments, the final two models with learned coefficients achieved an average 79% accuracy on fitting and predicting learners' learning outcomes of targeted grammatical features of applying CATE.

Figure 6b plotted the interactive efficacy among effort, salience and intervention rounds of acquiring possessive pronouns ($0.97 * \text{Exp}(E_{f1}, E_{f2}, S_r, E_2)$); and the interactive efficacy among effort, salience and intervention rounds of acquiring 'be' question features ($0.13 * \text{Exp}(E_{f1}, E_{f2}, S_r, E_2)$). It needs to be noted that all the results were the outputs of the trained models, therefore, there is a difference between a_k s for the first 3 rounds in Fig. 6c and a_k s in the Fig. 5. The outputs regarding epoch 4, 5, ...10, were estimated outputs of two trained models by setting parameter E_2 to 4, 5, ...0.10. It could be seen from Fig. 6b, c that the estimation of the coefficient b could be used to label the 'easy-perceiving level' of the target grammar for the participants, suggesting that the easier the target being perceived (small b), the quicker the ceiling of ESE's effects reach, and vice versa (large b). Figure 6b shows that, the effect of ESE regarding 'be'-question features reached the ceiling at the end of the first intervention round.

Figure 6c plotted the proportion correct α predicted by a computational model as a function of intervention rounds (E_2) with regard to G1, G2, G3 in possessive pronouns' acquisition, and the proportion correct α predicted by another computational model as a function of intervention rounds (E_2) with regard to G2 (content was presented in a batch mode) in 'be'-question features' acquisition. As shown, the proportion correct of G3 was expected to achieve 1 and the proportion correct of others converged in the range of [0.5, 0.65]. The serial curves suggested that under most conditions, the efficacy of CATE has a 'resistance trait.'

Discussion

The PESE computational model achieved prominent fitting and predicting results (achieved average 79% accuracy) on the data collected from two experiments, revealing a reasonable assumption of the efficacy of applying CATE in L2-grammatical-knowledge learning, by using learners' prior knowledge of the targets, their TE-dominant attention effort, the salience of the targets, and CATE-based intervention epochs as the modulating variables. The experiment results showed that the efficacy of CATE is not only a function of general prior knowledge (Han et al., 2008), but

also a function of the interactive effects of the other variables. The PESE model suggested that although prior knowledge of the target was the dominant factor of deciding the efficacy of the CATE-based intervention, the interactive effects of other variables could partly compensate for the effect discrepancies because of the lack of prior-knowledge. Moreover, the coefficient b , i.e., the easy-perceiving level of the target knowledge for learners decided the effect ceiling of repetition; whether the intervention had a ‘resistance trait’ or not was decided by the inductive process that involved with adequate information.

It has been widely accepted that the learners’ different prior knowledge may influence the efficacy of CATE, e.g., see Park and Warschauer (2016) and Han et al. (2008), and has been suggested to control learners’ prior knowledge before using focused CATE technologies (Meurers et al., 2010). This study had carefully controlled for the learners’ prior-knowledge differences by using SVO structure as the basic form of the sentences, and involving the words that had been sufficiently familiarised by learners in the intervention. However, the initial discrepancies of learners’ prior-knowledge among different groups showed significant influence on the intervention effect. It could be seen from Fig. 5 that, the students in G1 who had intermediate prior-knowledge of the targets achieved best intervention results at the end of the third epoch, which is in line with the results found by Park et al. (2012). Nevertheless, Fig. 6c illustrated that the interactive effects of ESE could partly compensate for this less-efficacy as the convergent proportion correct of G2 was superior to that of G1; and the positive L1-L2 knowledge-transfer quickly facilitated learners in conducting comprehension activities, which in consequence compensated for learners’ inadequate prior-knowledge of targets. The effect of L1-information was simulated through variable S in PESE model.

PESE model showed that the positive L1-L2 knowledge-transfer initiated a new inductive process, which consequently led to the significant improvements of the learners’ awareness of the target knowledge (see Fig. 5). In contrast, the efficacy of the interventions without L1-information quickly reached peaks and entered into a resistance period, see Fig. 6c. This ‘resistance trait’ of the CATE-based intervention had also been reported by previous studies, e.g., in Rott’s study (2007), learners achieved best performance just through one time of intervention, and in Szudarski and Carter (2016), their participants’ best performance appeared in the 6th epoch. Figure 6b showed that the increment of proportion correct made by ESE in each epoch was decided by the ‘easy-perceiving level’ of the grammar, i.e., the best intervention rounds could be predicated through coefficient b , and the coefficient b also can be estimated through experience data or by teachers. In this study, the proportion correct of ‘be’-question features converged at the third epoch, where $b=0.13$; in contrast, the proportion correct of possessive-pronoun features converged at about 5th epoch, where $b=0.97$. Moreover, PESE could also simulate the inductive efforts that override the effects of repetition (Eckerth & Tavakoli, 2012), it can be found in Fig. 6c that the inductive process initiated by the L1-L2 knowledge-transfer expedited the convergence speed of the proportion correct by converging at the 5th epoch, in contrast, the proportion correct of G2 was converged at the 6th epoch.

It is noteworthy that, the inductive efforts were not only appeared in G3, but also appeared in G1 and G2 as the attention effort invoked by learners was spent both

on targeted- and un-targeted area, suggesting that those grade-2 students actively behaved motivational-cognitive-involvement (Laufer & Hulstijn, 2001) in understanding the transformation of the subject pronouns to possessive pronouns. CATE could not prevent them from searching related information when they conducted this comprehension process (Table 2 revealed that the attention spent on non-targeted area was almost the same as the attention spent on the targeted area). Although Eckerth and Tavakoli (2012) demonstrated that the input–output cycle which was superior to the repeated input could be implemented by integrating the inductive activities into the reading tasks, PESE complemented that only the spontaneously behaved motivational-cognitive-involvement was not strong enough to help learners break the ‘resistance trait,’ unless the information that could initiate new inductive activities had also been involved in the cycle, e.g., the L1-information that could facilitate learners’ inductive activities of understanding the target knowledge.

Pedagogical implications

Based on the conclusions mentioned above, three pedagogical implications are proposed for English teachers who may use CATE in their classrooms at lower grades. First, a list of the ‘easy-perceiving levels’ of the potential knowledge that may be targeted at by using CATE should be built. It is recommended to capture the coefficient bs of the knowledge by fitting the data collected from empirical experiments into PESE model; or, alternatively, teachers could estimate the bs based on their experience by comparing the knowledge with those referred to in this study.

Second, the best repetition rounds of applying CATE for young EFL learners in L2-knowledge acquisition should be decided according to the ‘easy-perceiving level’ of the target (but not recommend to exceed 6 rounds). After reaching the efficacy ceiling of repetition, new information or instructions should be added into the intervention to initiate a new inductive process of knowledge comprehension, i.e., to change the status of variable S .

Third, despite the learners’ prior knowledge at intermediate level is the optimal prerequisite of applying CATE-based intervention, learners with insufficient prior knowledge could also benefit from the CATE-based intervention, i.e., the efficacy of the CATE-based intervention could be compensated for by strengthening the interactive effects of ESE. For example, by opportunely increasing the repetition number, by adding new information to stimulate further inductive activities, and by improving learners’ attention distribution if the attention they devoted into tasks was not satisfying.

Limitations and future work

This study employed a transparent linear model to simulate the interaction among prior knowledge, CATE-based visual attention, and repetition under different content modes in which CATE took place. The model provided answers for EFL teachers regarding questions like, why the effectiveness of applying CATE in classroom

is unstable, or to what extent that the interactive effects of ESE can compensate for the lack of the prior knowledge according to the ‘easy-perceiving level’ of the target knowledge. However, the current model remains the room for improvement in accuracy. Therefore, one of our future works will focus on studying the possibility of employing deep learning techniques with rule extraction function as the basic infrastructure for building model’s ESE part. The model is expected to be improved both in its accuracy in fitting and predicting the data collected from classroom experiments, and in its explanatory ability. Secondly, although the *bs* learned from the experimental data revealed their correspondence with the difficulty levels in learning targeted grammatical items, the mapping of the two remains in a rough relation which may prevent teachers from accurately estimating *b*. Providing a precisely correspondence between coefficients and the difficulty levels in learning grammatical items is our another ongoing works.

1. The sentences used in card-combination game were not be initiated with capital letters given that the participants were less familiar with the capital letters in contrast to lower-case letters.
2. The details of how the experiment conducted can be found in another manuscript of us which has been submitted to another journal and is currently under review.

Appendix A

Sentences used in the experiment.

Sentences		Experimental group		
		G1	G2	G3
Personal pronoun as subject	Possessive pronoun as subject			
He eats a cake	His brother eats a cake	✓	✓	✓
He has a pen	His mom has a pen	✓	✓	
He is a boy	His friend is a boy	✓	✓	✓
He is tall	His Dad is tall	✓	✓	✓
He has a dog	His Dad has dog			
It has a fish	Its friend has a fish	✓	✓	✓
It is floppy	Its name is floppy	✓	✓	✓
It is strong	Its dad is strong	✓	✓	✓
It is too big	Its face is too big	✓	✓	✓
It is red	Its colour is red	✓	✓	✓
She tells a story	Her brother tells a story	✓	✓	✓
She is sleeping	Her dad is sleeping	✓	✓	
She has a dog	Her mom has a dog	✓	✓	
She has a cat	Her sister has a cat	✓	✓	✓
She is a girl	Her friend is a girl			✓
She eats a cake	Her Mom eats a cake			✓

Acknowledgements We would like to express our thanks to all people who have ever helped us in this paper. First of all, we would like to make thanks to the reviewers and editors who helped us a lot in improving the quality of the paper; Then, we would like to make thanks to the Director of Developmental Neurocognition Lab, Professor Michael S. C. Thomas, who gave us many advises on building of the computational models. Funding was provided by Natural Science Foundation of Sichuan Province (Grant No. 2018JY0312) and Social Science Foundation of Sichuan Province (Grant No. 2021A022).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alanen, R. (1995). *Input enhancement and rule presentation in second language acquisition Attention and awareness in foreign language learning* (pp. 259–302). University of Hawaii press.
- Aloysius, N., & Geetha, M. (2017). *A review on deep convolutional neural networks*. Paper presented at the International Conference on Communication and Signal Processing (ICCSPP), pp. 588–592.
- Boers, F., Demecheleer, M., He, L., Deconinck, J., Stengers, H., & Eyckmans, J. (2017). Typographic enhancement of multiword units in second language text. *International Journal of Applied Linguistics*, 27(2), 448–469.
- Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1), 185–207.
- Borji, A., Sihite, D. N., & Itti, L. (2011). *Computational modeling of top-down visual attention in interactive environments*. Paper presented at the 22nd British Machine Vision Conference, Scotland, UK, pp. 85.1–85.12.
- Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Saliency in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology*, 7, 1284.
- Comeaux, I., & McDonald, J. L. (2018). Determining the effectiveness of visual input enhancement across multiple linguistic cues. *Language Learning*, 68(1), 5–45.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. Paper presented at the Proceedings of the 2019 Conference of the North American (Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1), pp. 4171–4186.
- Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekke, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 22(4), 657–679.
- Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research*, 16(2), 227–252.
- Elgort, I. (2017). Technology-mediated second language vocabulary development: A review of trends in research methodology. *CALICO Journal*, 35(1), 1–29.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
- Ellis, N., & Sagarra, N. (2010). Learned Attention Effects in L2 Temporal Reference: The First Hour and the Next Eight Semesters. *Language Learning*, 60(2), 85–108.
- Ellis, N., & Sagarra, N. (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in second language acquisition*, 33(4), 589–624.
- Gascoigne, C. (2006). Explicit input enhancement: Effects on target and non-target aspects of second language acquisition. *Foreign Language Annals*, 39(4), 551–564.

- Gollan, B., & Ferscha, A. (2016). *SEEV-effort—Is it enough to model human attentional behavior in public display settings*. Paper presented at the Eighth International Conference on Future Computational Technologies and Applications (FUTURE COMPUTING 2016), Rome, Italy, pp. 8–14.
- Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics*, 29(4), 597–618.
- Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied*, 12(2), 67–78.
- Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions: An eye-tracking study. *Studies in second language acquisition*, 39(3), 401–430.
- Kim, Y. (2006). Effects of input elaboration on vocabulary acquisition through reading by Korean learners of English as a foreign language. *Tesol Quarterly*, 40(2), 341–373.
- LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research*, 20(1), 75–91.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, 57(1), 87–118.
- Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition*, 30(3), 307–331.
- Lin, P. M. (2014). Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System*, 42, 164–176.
- Liu, Y. T., & Leveridge, A. N. (2017). E-books and implicit L2 vocabulary acquisition. *British Journal of Educational Technology*, 48, 43–56.
- Mareschal, D., & Thomas, M. S. C. (2007). Computational modeling in developmental psychology. *IEEE Transactions on Evolutionary Computation*, 11(2), 137–150.
- Mayén, N. R. (2013). Effects of input enhancement and visual prompts in children's L2 acquisition of Spanish Verbal morphology. *ELIA: Estudios de Lingüística Inglesa Aplicada*, 13, 83–111.
- McGowan, V. A., & Reichle, E. D. (2018). The “risky” reading strategy revisited: New simulations using E-Z Reader. *Quarterly Journal of Experimental Psychology*, 71(1), 179–189.
- Meurers, D., Ziai, R., Amaral, L., Boyd, A., Dimitrov, A., Metcalf, V., & Ott, N. (2010). *Enhancing authentic web pages for language learners*. Paper presented at the Fifth Workshop on Innovative Use of NLP for Building Educational Applications, Los Angeles, pp. 10–18.
- Monaghan, P. J., Chang, Y.-N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language*, 93, 1–21.
- Park, H., Choi, S., & Lee, M. (2012). Visual input enhancement, attention, grammar learning, & reading comprehension: An eye movement study. *English Teaching*, 67(4), 241–265.
- Park, Y., & Warschauer, M. (2016). Syntactic enhancement and second language literacy: An experimental study. *Language Learning & Technology*, 20(3), 180–199.
- Rast, R. (2008). *Foreign language input: Initial processing*. Clevedon, UK: Multilingual Matters.
- Reichle, E. D. (2015). Computational models of reading: A primer. *Language and Linguistics Compass*, 9(7), 271–284.
- Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, 57(2), 165–199.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Schmidtke, D., & Moro, A. L. (2021). Determinants of word-reading development in English learner university students: A longitudinal eye movement study. *Reading Research Quarterly*, 56(4), 819–854.
- Smith, M. S. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15(2), 165–179.
- Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, 26(2), 245–265.
- Uliitsky, H. (2000). Language learner strategies with technology. *Journal of Educational Computing Research*, 22(3), 285–322.
- Walker, S., Schloss, P., Fletcher, C. R., Vogel, C. A., & Walker, R. (2005). Visual-syntactic text formatting: A new method to enhance online reading. *Reading Online*, 8(6), 1–27.

- Wang, X., Zhao, X., & Ren, J. (2019). A new type of eye movement model based on recurrent neural networks for simulating the gaze behavior of human reading. *Complexity*, 2019, 1–12.
- Wang, X., Zhao, X., & Xia, M. (2018). *The prediction model of saccade target based on LSTM-CRF for Chinese reading*. Paper presented at the International Conference on Brain Inspired Cognitive Systems, pp. 44–53.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91–120.
- Wickens, C. D., Goh, J., Helleburg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multi-task pilot performance using advanced display technology. *Human Factors*, (45), 360–380.
- Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition*, 35(2), 323–352.
- Yang, J., Qi, X. F., Wang, L., Sun, B., & Zheng, M. X. (2022). A reading model of young EFL learners regarding attention, cognitive-load and auditory-assistance. *The Journal of Educational Research*, 115(1), 51–63.
- Yang, J., Thomas, M. S. C., Qi, X. F., & Liu, X. (2019). Using an ANN-based computational model to simulate and evaluate Chinese students' individualized cognitive abilities important in their English acquisition. *Computer Assisted Language Learning*, 32(4), 366–397.
- Yusuf, A. B., Kor, A., & Tawfik, H. (2019). *Development of a simulation experiment to investigate in-flight startle using fuzzy cognitive maps and pupillometry*. Paper presented at the international joint conference on neural network, Macao, China, pp. 1–10.
- Ziegler, N., Meurers, D., Rebuschat, P., Ruiz, S., Moreno-Vega, J. L., Chinkina, M., Li, W., & Grey, S. (2017). Interdisciplinary research at the intersection of CALL, NLP, and SLA: Methodological implications from an input enhancement project. *Language Learning*, 67(S1), 209–231.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.