

## An Adaptive English Video-recommendation Learning System Based on Latent Semantic Analysis and Non-mastery Conceptions

Pei-Ling Tsai<sup>1</sup> Cheng-Hsuan Li<sup>2</sup> Shih-Yen Hung<sup>3</sup>

According to technology development, more and more English video resources can be used to recommend to learners to learn English. However, among the vast video sources, how to give appropriate teaching videos according to the learners' abilities is a critical research topic. This study proposes to use latent semantic analysis integrated with students' non-mastery concepts to recommend appropriate learning videos to improve students' non-mastery conceptions. Video recommendation corpus texts were analyzed using term weights and Latent Semantic Analysis (LSA) rules. Based on learners' error responses, enhanced English instructions were developed for the video recommendation system. A total of 48 first-year students participated in the experiment. Students in the experimental group learned English with recommended videos through the proposed recommendation system, and students in the control group learned English according to the list of videos. The experimental results found that recommended English learning videos based on learners' non-mastery concepts can improve learners' English listening and reading levels. The learning effect is better than that of the control group.

**Keywords:** adaptive learning, English as a Foreign Language, media in education, recommendation system

1. Pei-Ling Tsai, Adjunct Assistant Professor, Center for General Education, National Taichung University of Education
2. Cheng-Hsuan Li, Professor, Graduate Institute of Educational Information and Measurement, National Taichung University of Education
3. Shih-Yen Hung, Master, Graduate Institute of Educational Information and Measurement, National Taichung University of Education

Received: 2022.03.15; Revised: 2022.11.04; Accepted: 2023.02.19

Corresponding Author: Cheng-Hsuan Li; Email: chenghsuanli@gmail.com

Address: No. 140, Minsheng Rd., West Dist., Taichung City 403514, Taiwan

Graduate Institute of Educational Information and Measurement, National Taichung University of Education

## 基於潛在語義分析和非精熟概念之適性英語影片推薦學習系統

蔡佩玲<sup>1</sup> 李政軒<sup>2</sup> 洪士彥<sup>3</sup>

### 摘要

在網路發達的年代，有許多英語影片資源可用來推薦給學習者進行英語學習。然而，在茫茫的影片來源中，如何根據學習者的能力給予適當的教學影片，則是一個很重要的研究課題。本研究提出利用潛在語義分析技術，比對學生不精熟概念與影片字幕來推薦適當的學習影片，補強學生不精熟概念。利用語詞權重和潛在語義分析 (Latent Semantic Analysis, LSA) 影片字幕 (texts)，再根據學生不精熟的概念推薦學習影片。共有 48 位大一新生參與實驗，實驗組學生將會透過推薦系統的推薦影片進行學習，對照組則根據影片列表來進行學習。實驗結果發現，基於學習者不精熟概念推薦英語學習影片可以提升學生的英語聽力和閱讀水平以及學習成績，學習成效也優於對照組。

**關鍵詞：**英語學習、推薦系統、教育媒體、適性學習

- 
1. 蔡佩玲，國立臺中教育大學通識教育中心兼任助理教授
  2. 李政軒，國立臺中教育大學教育資訊與測驗統計研究所教授
  3. 洪士彥，國立臺中教育大學教育資訊與測驗統計研究所碩士生
- 收件日期：2022.03.15；完成修改：2022.11.04；正式接受：2023.02.19  
通訊作者：李政軒；Email：chenghsuanli@gmail.com  
地址：403514 臺中市西區民生路 140 號  
國立臺中教育大學教育資訊與測驗統計研究所

## Introduction

“English as foreign language learning requires the support of various learning tools to offer additional opportunities to learn it” (Chen & Li, 2010, p. 432). The higher the advancement in technology, the more dynamic, complex, and interactive is the learning methodology and the instructions provided for learning English (Cardoso, 2011; Conole, 2012;; Lan et al., 2007; Laurillard, 2013; Lincoln, 2011; Luckin, 2010; Roschelle et al., 2013; Sun, 2018). Computer technology has reinforced learning effectiveness (Altınay-Gazi & Altınay-Aksal, 2017; Chen et al., 2018) and has lowered the anxiety caused while learning languages (Lee et al., 2017). Although various resources for learning English through videos have been created in rapid succession, they recommend learning materials based on personal interests and the interactions taking place between the browser and the system server. Through these media and systems such as BBC Learning English, Youglish, and LLN: Language Learning with Netflix, people learn English for convenience and autonomy. These systems take advantage of the subtitle function so that the learners grasp the meaning and content of the whole video. Without making a distinction between the learning materials and targets, users might withdraw their learning from the incomprehension of advanced content. On the contrary, learning basic content in place of challenges lowers the interests.

A number of studies applied personalized recommendation systems to online learning based on learning ability and learner’s major subjects (Baker, 2001; Chen, 2008; Chen & Chung, 2008; Chen et al., 2005; Hsu et al., 2013; Shu et al., 2018). In the study by Hsu in 2008, an English learning recommender system based on reading material was designed using two data mining techniques. Every English lesson had a basic score, and students obtained a higher score for matching their own interests with the recommended lesson after each lesson. The system recalculated the score of each student and recommended another reading lesson or material. The English lesson with the recommender system worked effectively (Hsu, 2008). Many studies developed systems that recommend English learners where and what to begin learning are mostly based on teaching texts or learners’ subjective interest (Amiri & Branch,

2012; Chen & Li, 2010; Wang & Liao, 2011). Other systems utilize search engines to find proper learning materials based on learners' ability on major subjects (Chen et al., 2005). Due to the accessibility, English instruction media have been commonly utilized to improve the instructions and enhance the learning effectiveness in a school environment (Caldwell, 2007; Chen & Li, 2010; Lee et al., 2017). Moreover, a system with personal error analysis is necessary which helps students to learn based on their errors and simultaneously allows teachers to monitor students' time of learning and their responses for further teaching. However, few studies have provided recommendations based on learners' error responses, nor systems were designed for learning from learners' errors. Referring to Gass (2013) researched "error analysis" and found that learners' error was regarded as the signals of their learning system in terms of the subconscious knowledge in the target language. The significance of language learners' error represents the process of learners attempting to comprehend the target language and infer the usage regulation. Instructors tell learners mastery of the target language from the errors so as to assist in proceeding learning. Additionally, some systems adopt Latent Semantic Analysis (LSA) (Martin & Berry, 2007) as the primary method for recommendation and the most relevant texts were recommended to a learner after matching them with the system corpus (Forsyth et al., 2018; Lee, 2019; Lee et al., 2013; Lee et al., 2009). The texts of video included in the corpus of the video recommendation system were further transformed as a term-by-document matrix and analyzed using the term weights and based on rules of LSA. Hence, a video recommendation system that is based on learners' error responses and provides enhanced English instructions was designed with the aforementioned best practices in mind. This study proposed an adaptive English learning method through a video recommendation system based on learners' errors analysis by LSA. The questions that guided this study are presented as follows:

1. Do the English instructions provided with video recommendations based on learner's error analysis improve English listening and reading proficiency?
2. Is the adaptive English learning system based on video recommendation more effective than the one without?

This study proposes to learn English through videos and scripts. The personalized English instruction video recommendation system recommends to learners where

and what to begin learning based on the similarities between scripts of videos and the learners' error responses.

## Literature review

Audiovisual aids help language learners to have a general picture of the target language, to extensively increase their access to visual and auditory communication, and to motivate them while learning (Chen, 2011; Hsu, 2018; Tschirner, 2001; Yang, 2011). Using multimedia as a learning aid improves learners' concentration and motivation. Some studies established computerized adaptive recommendation systems or web-based recommendation systems based on theories, like Item Response Theory, and took learning content and learner's ability into account in order to improve English ability and learning performance (Baker, 2001; Chen, 2008; Chen & Chung, 2008; Chen et al., 2005; Hsu et al., 2013; Shu et al., 2018). Moreover, the audio and visual stimulations place language learners in an authentic environment that encourages the learner to use the target language to communicate. Moreover, through multimedia, observing how the language is constructed is important to learners and realistic settings are provided. Language learners participating in a class activity can encounter many real issues. The English instruction supplemented by the videos conduces learning effectiveness and learning interests (Hung & Chen, 2018; Lan et al., 2007). Although non-native English speakers are not accustomed to speaking, listening, or reading English, they can read video scripts. These scripts are nearly similar to other types of textual materials but are more lively and entertaining (Tschirner, 2001). As Stockwell et al. (1965, p. 5) described, "Grammar is a description of the probability relationships between lexical units in sequence," reading video texts is similar to learning and reading articles in the target language. However, the texts of videos comprise several short sentences, and the scene and sound assist learners to comprehend the meaning and usage of the English vocabulary, slang, and phrases used in the context. The English instructions trigger learners' interest in practicing due to the acousto-optic effect (sound and light effect). Furthermore, the learners' interest arises from the English-speaking content in a video with sound and light effects (Hung & Chen, 2018).

This study adopts two frameworks—the Lexile analyser and Latent (s)emantic (a)nalysis (LSA). First, based on the Lexile analyzer and common European framework of reference for languages (CEFR)—to scale videos according to the scripts and recommend videos to learners based on their error responses and analysis by using the video recommendation system. The CEFR was developed by the Council of Europe for evaluating and equalizing different language examinations (Figueras & Noijons, 2009; North, 2007) and several language examinations, online language systems, and some studies have adopted CEFR (Lawley, 2016; Velleman & van der Geest, 2014) to clarify different materials as well as the language testing criteria. Language examinations are linked to the CEFR as a diagnostic result is linked to a patient. Moreover, the error responses of the examinations assist learners to determine where they should begin reading or learning and which level of examination they should take next (Figueras & Noijons, 2009; Velleman & van der Geest, 2014). In analyzing ‘the role of the natural language,’ Gass (2013) proposes that learners’ error analysis helps teachers define learners’ perception and knowledge during the cognition process of learning the target language. Six steps of the error analysis are relatively significant for both learners and teachers to collect data, identify errors, classify errors, quantify errors, analyze the source, and remediate. Learners infer the meaning of the target language based on native language so that teachers can conclude the possible reason behind the grammatical or comprehensive errors in order to break through the learning myth which is the errors learners made during recognizing the usage of the target language and to encourage the learning progress. The Lexile analyzer, an online text analyzer, was established by MetaMetrics. The analyzer system uses an algorithm to classify texts, break down texts, analyse terms in the texts and then search for the matching text; such as paragraphs, articles, books, and other reading material (Lennon & Burdick, 2004; Smith et al. 2016; Smith & Turner, 2016). Moreover, a similar analyzer was adopted as a basis in the reference database and helped to note the errors made by nonnative speakers (Futagi et al., 2008). In this study, the video scripts were analyzed using the Lexile analyzer as vocabulary analysis to match with learner’s errors analysis and making a decision to use appropriate accessorial teaching materials as the base for video recommendation.

Based on Landauer et al. (1998, p. 2), “Latent (s)emantic (a)nalysis (LSA) is a

theory and method for extracting and representing the contextual usage meaning of words by statistical computations applied to a large corpus of text” (Landauer et al., 1998). The LSA has been applied meritoriously to many studies in fields such as reading, music therapy, medication, language learning, and education (Forsyth et al., 2018; Kuo et al., 2018; Landauer & Psotha, 2000; Lee, 2019; Lee et al., 2013; Lee et al., 2009). These systems adopted LSA as the major method and then gave feedback after matching a learner’s text with the system corpus. In the study by Forsyth et al. in 2018, a conversation-based system that was developed by natural language processing and LSA assessed learners’ proficiency and investigated the learning progress. The results showed that learners had improved higher word count and more complex language usage comparing person to person interviews and assessments. The study concluded that the system could evaluate learners’ proficiency and moreover, applying the system based on LSA assessment was achievable. Moreover, in the study by Lee et al. in 2013, an essay critiquing system (ECS) applying LSA was used to give feedback to learners’ writing. In Lee’s study in 2019, a system could detect learners’ writing texts and give instant feedback by LSA, and in the study of Lee et al. in 2009, a web-based essay critiquing system was developed by LSA to give feedback on learners’ writing. Additionally, in the study of Sung et al. in 2016, a recommended system matching learner’s summary with expert’s assessment based on LSA gave scores and feedback to learners and it concluded the proposed system enhanced learner’s writing.

Thus, this study developed not only the LSA in the video recommendation system to calculate the term weights which represent every meaningful letter in the video scripts but also analyzes learners’ error responses in the assignment by the developed system. Furthermore, the system gives a recommendation of a video from the matching words, terms, and clauses in learners’ error responses. Learners have chances to look at the usage of vocabulary, clauses, and sentences with which they are not familiar again for the purpose of reviewing the misunderstandings. Tinkham’s study in 1993 concluded that the semantic clusters like whisk, strainer, spatula... etc., had an effect on language learning, especially vocabulary (Tinkham, 1993). Therefore, LSA can also measure the similarity between source texts and learners’ error responses and this research was to learn the effect of the video recommendation system based on

learner's error analysis by LSA.

## Developing the video recommendation system

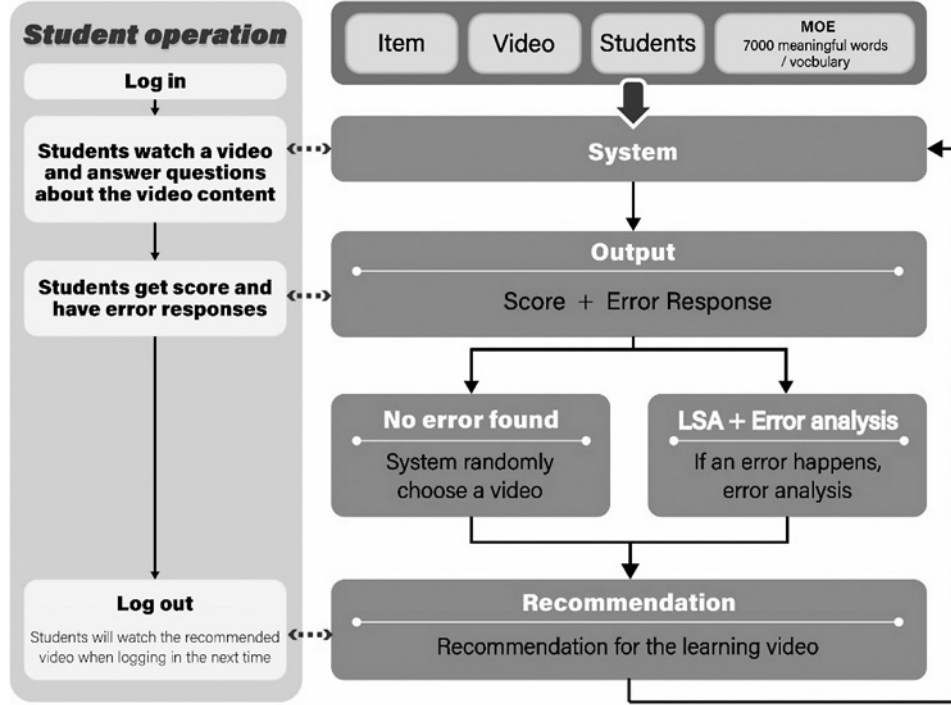
The video recommendation system is built on an online platform using PHP. The system is simple but has a practical recommendation system platform so that students can use the online system easily for learning (Lai et al., 2016). First, students enter the system, watch a video picked randomly by the system, and answer five questions that have multiple choices. Every student receives the same initial video and test at the first entry. If students answer five questions correctly at the previous stage, the system randomly collects a video for their next learning material. However, if a video was recommended earlier, they watched a recommended video based on the record of the previous responses. Learners logged into the system and watched a video once every two weeks.

Their responses to the questions related to the video were recorded and especially the error responses were analyzed using the video recommendation system. The system suggests the next video to students by using the LSA and searching for the best similarities or the closest weight values between the meaningful words of questions responded wrongly and the scripts of remaining videos for the next time they log in the system. The system using LSA analyzes learners' error responses to find out all the meaningful words in the responses and questions. Then by matching the meaningful words with the text of videos, the best recommendation will be made to learners.

Figure 1 displays the framework and learner operation process of the video recommendation system. Four databases construct the video recommendation system and learners watch a video and answer questions after logging into the system. If a student selects error distractors to a question, it is evident that the student has misunderstood the video text. The video subtitles or script is another type of readable material (Hung & Chen, 2018). The error responses in the term space are then analyzed using the personalized English learning video recommendation system that employs text mining measurement and an LSA. The formula below is the operation and recommendation rule of LSA. The  $w$  in the formula represents the texts or words that stu-



**Figure 1.** The system framework corresponding to the learner operation procedure.



dents cannot understand. Then, the similarities between the error responses and a video script are matched using the following rule:

$$\max_i \left\{ \frac{w^T v_i}{\|w\| \|v_i\|} \mid i = 1, 2, \dots, s \right\} \quad (2)$$

Subsequently, the system recommends the next video based on the learner's misunderstanding of the English text. By conducting the repeated review, the participant may enhance their ability to comprehend spoken English and read English text.

### Video and semantic space

In Kennedy and Miceli's research in 2017, students learn writing patterns from the recommended terms in the corpus with an open search method, which enables them to use and learn the language more accurately. In order to build a corpus, a lexical-file matrix was established first, which is a two-dimensional spatial matrix with

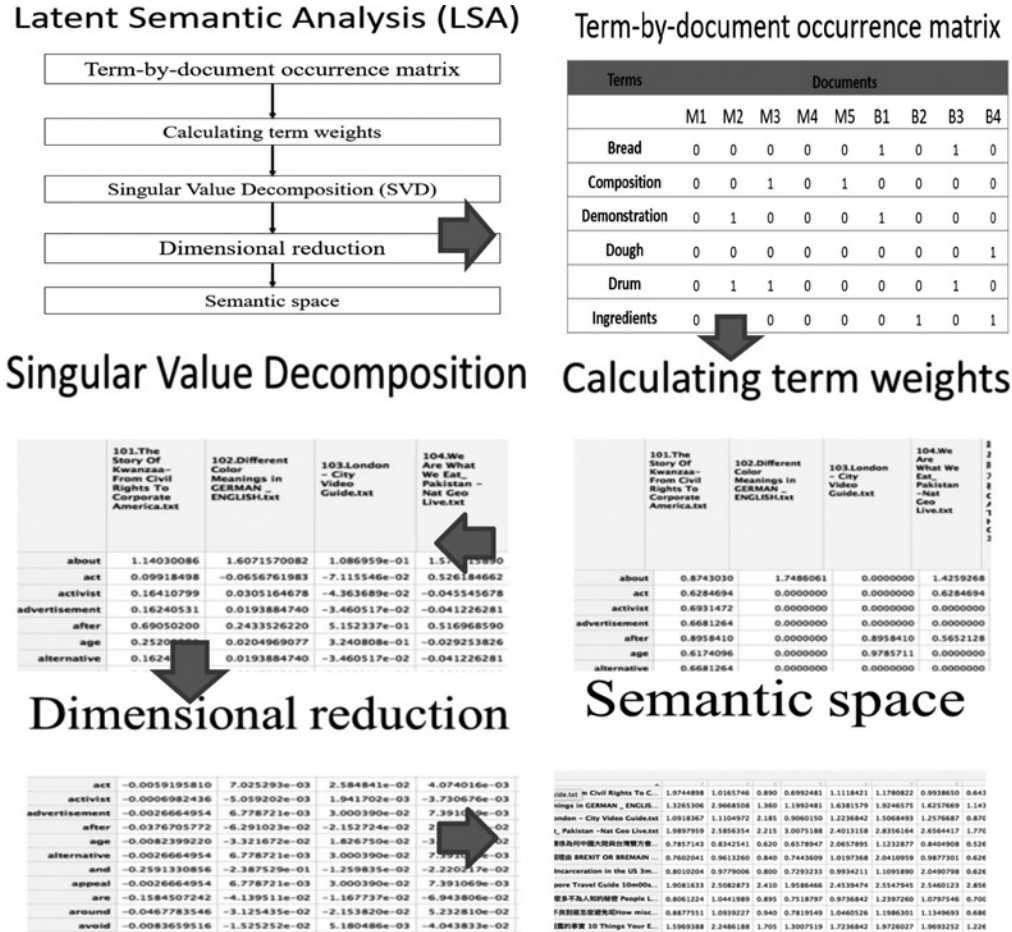
terms as columns and documents or articles as rows, where these selected terms are non-repeating terms defined by the corpus. 7,000 English vocabulary listed by Taiwan College Entrance Examination Center was as the basic terms. Furthermore, approximately 100 more online English-speaking videos were collected by experts and analyzed by the Lexile analyzer to verify the difficulty level of videos and specify level B of the CEFR and 90 videos were selected. This research adopted the Lexile analyzer and then the experts as filtering mechanisms based on the criteria containing the length of the video text, video text complexity, oral speed, and accent. It is necessary that the language learners implement existing knowledge in their native language to conduct a foreign language-based cultural context (Neville et al., 2009). The subjects of videos were related to diverse cultures and the teaching and learning videos obtained from YouTube that were uploaded by people worldwide because learning English through different cultures helps learners to notice the English usage by various English speakers and how to use it properly. The lengths of the selected videos are within 10 minutes for a few main concepts in a video to maintain the motivation and interests of learning English.

According to Martin & Berry, 2007, the texts of the 90 videos included in the corpus of the video recommendation system were further transformed as a term-by-document matrix and analyzed using the term weights and based on rules of LSA as well as the “tm” suite in the R system to remove punctuation marks and numbers from texts, changing uppercase letters to lowercase letters, and extracting stems.

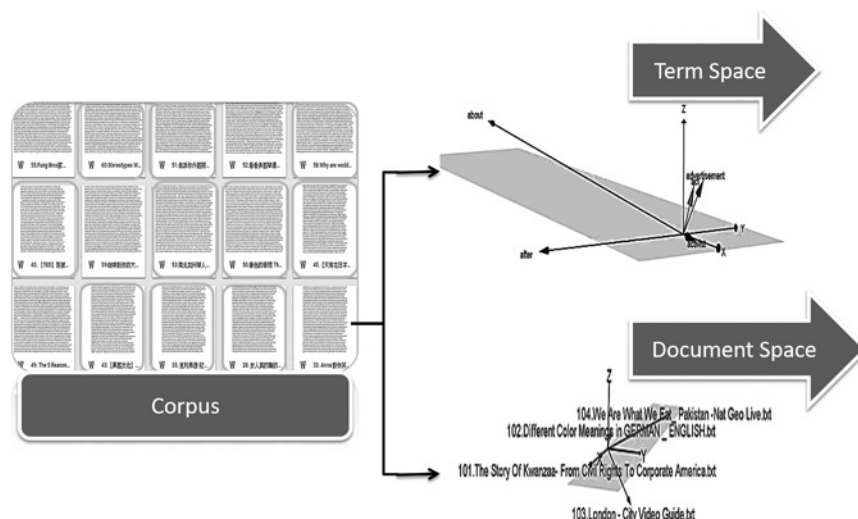
LSA technique refers to a long list of words or content, from which to analyze the summary and the general idea of the paragraph, or even further, to organize the meaning of the whole article. Every word or term has its weight representing how important and often the word shows up in the article. The weight of meaningful words was increased in the sentences and then reduced or even deleted the reduplicative words that appear in every text, such as ‘a,’ ‘the,’ ‘or,’ etc. Moreover, it was decomposed with the singular values and reduced the dimension. The singular value analysis (SVD) method is to calculate the eigenvalue of each keyword in the diagonal matrix after the operation. Each word is no longer just the frequency of occurrence, but a more representative meaning. That is to say, the larger the eigenvalue of the vector, the larger its information volume. The flow chart of the LSA basic structure is shown

in Figure 2. As a result, every video text has turned into a new and representative semantic space and the corpus of the video recommendation system is divided into two spaces, as presented in the following Figure 3.

**Figure 2.** The flow chart of the LSA basic structure.



**Figure 3.** *Corpus and the semantic space of the recommendation system.*



### Video recommendation system

Before starting to build the recommendation system, the pre-set propositions and analysis questions related to the video content were completed by the experts and the tm suit of the R language system, and then the completed analysis content was set in the system. The recommendation system is built on the network platform by means of HTML and PHP to write dynamic web pages so that students can watch movies and answer questions on the platform, and record all responses at the same time, and then recommend new videos. The analysis questions are based on the main concepts in the video such as the video “Things Asian girls hate” introducing how Asian American girls are different from Asians and Americans and their cultural background. Based on their cultural differences and upbringing, the video discussed things like western standards of beauty, hard-core music, etc. Referring to Bowyer-Crane and Snowling’s study in 2005, there are six types of comprehensive inferences that help teachers classify learners’ understanding difficulties and the experts of this study built up five reading comprehension questions with four-choice items, corresponding to every video. Accordingly, the questions and the corresponding types of comprehension for the video “Things Asian girls hate” were listed in Table 1.

**Table 1.** *Questions and the corresponding types of comprehension.*

Questions	Relation to the video	Types of comprehension
Why do Asian girls hate the sun?	Background knowledge	literal information and knowledge-based inferences
What does parallel parking mean?	Vocabulary	knowledge-based inferences and vocabulary dependent
What is the thing Asian girls hate?	Main concept	literal information and knowledge-based inferences
I'm versatile! I got options. What does "versatile" mean?	Vocabulary and video text comprehension	vocabulary dependent and cohesive inferences
According to the video, what is PAMS?	Term and video text comprehension	literal information and vocabulary dependent

For example, when learners answer some of the questions for the video "Things Asian girls hate" incorrectly, the next video "15 SURPRISING CULTURAL DIFFERENCES to a Japanese in America" related to "Asian, girl, sun, like, care, etc.," the keywords of questions learners answered wrongly, will be recommended. The meaningful words in the question and the items are the marked terms in Figure 4, like Asian, girl, hate, sun, like, tan, etc. The question "why do Asian girls hate the sun" indicates one of the crucial concepts in the video. The four red words in the question are the meaningful words and the system gives a recommendation of a video by matching the meaningful words in learners' error responses.

When learners answer questions, they will review the main concepts of a video and the semantic space was set consequently to about 7,000 English words based on the Taiwan vocabulary list. The meaningful words in the video script were represented as 7,000 different vectors, that is,  $t_1, t_2, \dots, t_{7000}$ . Additionally, the videos were considered as vectors  $v_1, v_2, \dots, v_s$ , where  $s$  represents the number of videos. Hence, each of the five questions with multiple choices that were related to a video was formed using the 7,000 words. In Figure 4, item  $i$  comprises the following terms: Asian, girl, hate, sun, like, tan, etc. The corresponding word vectors are and belong to. The vectors of the item  $i$  can be calculated using the following equation:

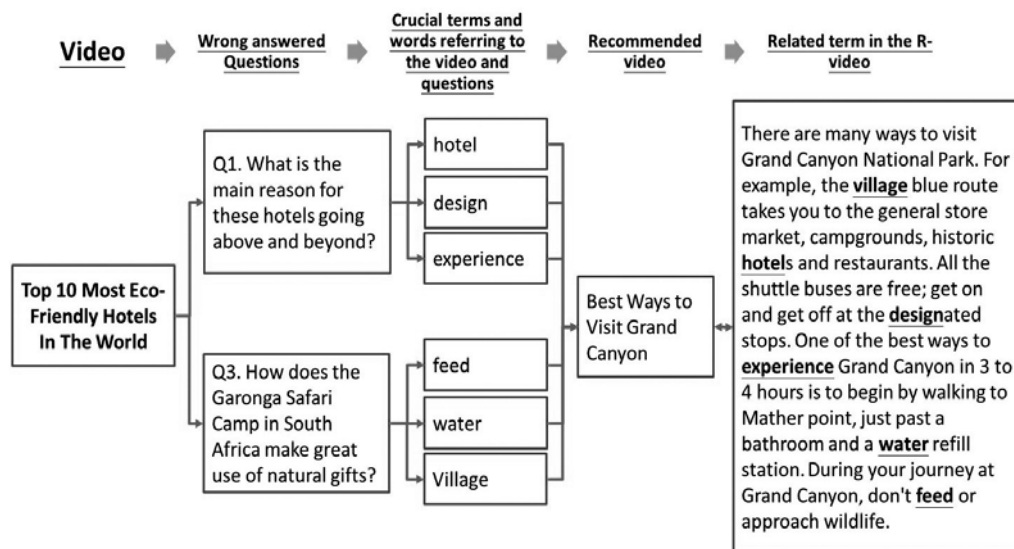
**Figure 4.** *Decomposition of the multiple choices of a question and its items .*

$$w_i = \frac{1}{14} \sum_{j=1}^{14} t_j^{(i)} \quad (1)$$

why do asian girls hate the sun  
 they like to be tanned  
 they care less about having white skin  
 they peek through the windows  
 they grow up that way

Through questions and answers related to the video text, the system records how learners comprehend the video and how they do not. For example as Figure 5, the video “Top 10 Most Eco-Friendly Hotels In The World” was the first video for student A. After watching the video, student A had answered two questions wrong out of five, and the crucial terms derived from these wrong answered questions and choice items were “hotel, design, experience, feed, water, and village. Therefore, the system recommended the next video “Best Ways to Visit Grand Canyon” for student A.

**Figure 5.** *Example of the recommending and analysing feedback.*



## Method

### Participants and measuring tool

48 students participated in the experiment from two freshman English course classes at the university in Taiwan. The participants had similar academic backgrounds such as studying English and taking many tests before the experiment for almost ten years before the study took place. There were 28 students assigned to the experimental group. These students were asked to log into the video recommendation system to watch a video and answer five questions as homework once in two weeks besides their regular course. The remaining 20 students were given the same instruction in class and print-based text with the worksheet as assignments; they formed the control group. Based on the test results of the pretest, all research participants belonged to the level between B1 and B2, intermediate proficiency in English; thus, the intermediate level was assigned using CEFR.

During the experiment, all the students took a pretest, a posttest and a delayed posttest which were from the General English Proficiency Test administered by the Language Training and Testing Center and supported by MOE in Taiwan. The pretest, posttest and delayed posttest were simulations of the General English Proficiency Test (GEPT) based on validity tests founded by the Language Training and Testing Center (LTTC) in Taiwan. The listening test questions and the reading questions are comprised of 45 and 40 multiple-choice items, separately. The perfect score is 240. Some studies have examined the correlation between learners' vocabulary size and the GEPT-test and found that their vocabulary size has a higher correlation with the achievement of the GEPT-test (Li & Huang, 2008; Shen & Wu, 2009). Moreover, there are researches regarding the validity of the tests, which are posted on the website of [https://www.lttc.ntu.edu.tw/e\\_lttc/e\\_gept/alignment.htm](https://www.lttc.ntu.edu.tw/e_lttc/e_gept/alignment.htm) according to the official website of LTTC. The test included two parts—listening and reading to inspect the students' ability to listen and read English effectively through the experimental process.

## Experimental Procedure

The experiment was conducted within 12 weeks and multicultural-related texts to be the teaching and learning material in the freshmen English courses. The instruction included reading articles, worksheets, and videos. A thematic text was taught every week in class to both groups and also a short video related to the text was watched along with the teaching. Both groups were taught by means of the English instructions based on the lesson plan. Furthermore, the control group was assigned with a print-based article and a worksheet as homework and the one for the experimental group was logging in the English video recommendation system as the treatment.

At the beginning of the experiment, the students took a pretest which was from the General English Proficiency Test. The control group was instructed with the same multicultural context and print-based text assignments while the experimental group was asked to use the video recommendation system for watching a video and answer five questions that had four multiple choices based on the text of the video they watched as homework once in two weeks besides the regular course. The error responses to the questions associated with listening, vocabulary, or reading comprehension were analyzed using LSA analysis in the video recommendation system. The similarity vector of words in both the error responses and the video scripts were calculated and matched to suggest the subsequent assignment video for participants. This procedure, as shown in Figure 6, was repeated every week for 12 weeks by the students in both groups. For verifying the experimental progress and the effectiveness of listening and reading ability separately after the pretest, both the groups were assessed using a posttest and a delayed posttest.

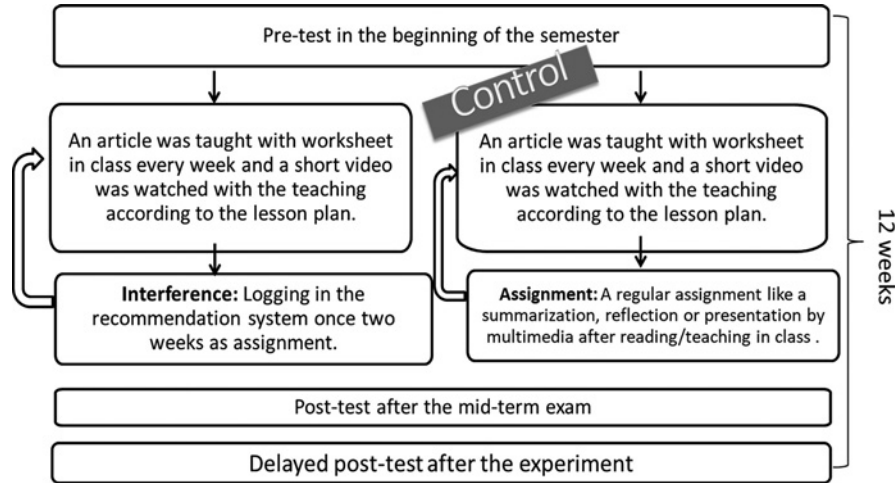
## Results and discussion

### Data analysis

To verify the experimental progress and effectiveness of listening and reading, the participants of both the groups took three GEPT simulation tests for the 12-week experiment—pretest, posttest, and delayed posttest. The *t-test* and one-way analysis



**Figure 6.** *The procedure of the experiment.*



of variance (ANOVA) are the two most common tests used for verifying data. The *t-test* is a statistical hypothesis test based on learners' distribution if the null hypothesis is supported. ANOVA was mainly used to compare whether there was a difference in the average values of the two groups. Analysis of covariance (ANCOVA) combined the features of ANOVA and regression analysis. The data analysis was mainly conducted using *t-test* and ANCOVA for detecting the primary and interaction effects.

20 participants in the control group (Group 1) were taught reading a print-based text and a worksheet as homework and 28 participants in the experimental group (Group 2) were requested to use the video recommendation system for 12 weeks. The means of the pretest scores for the control and experimental groups were 139.18 and 147.83, respectively. Moreover, the means of the posttest scores were 139.50 and 167.04, respectively. The means of the delayed posttest scores were 141.10 and 159.07, respectively, as shown in Table 2. The posttest scores obtained using the *t-test* for both the groups were found to be statistically significant ( $t = -3.065$ ,  $p = .004$ ,  $p < .01$ ). By analyzing the means of the scores presented in Figure 7, it was found that the pretest and posttest scores of the control group (Group 1) exhibited negligible difference (score: 139.18 and 139.50, respectively) during the experiment.

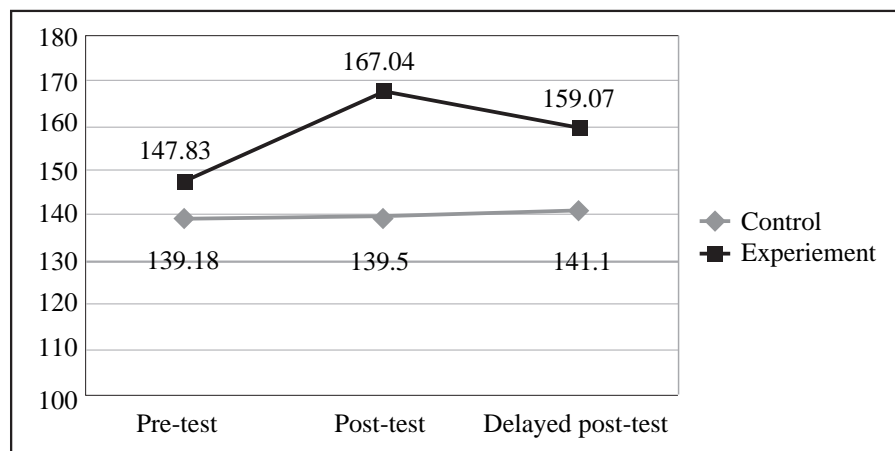
Moreover, the delayed posttest score exhibited a slight improvement from the pretest score (from 139.50 to 141.10). Due to the video recommendation system, the performance of the experimental group (Group 2) improved excessively (from 147.83 to 167.04) during the 12-week experimentation in which English instructions were provided with video recommendations based on a learner's error responses. However, the score degraded from 167.04 to 159.07 after the experiment when the system was not used.

**Table 2.** Scores of the three tests for the control group (Group 1) and experimental group (Group 2).

	Group	No	Mean	SD	<i>t</i>	Sig. (2-tailed)
Pre-test	1	20	139.18	39.450	-.846	.404
	2	28	147.83	27.342		
Post-test	1	20	139.50	36.986	-3.065	.004**
	2	28	167.04	25.338		
Delayed post-test	1	20	141.10	36.899	-1.773	.083
	2	28	159.07	32.907		

\*\* $p < .01$ .

**Figure 7.** Line chart of the three tests for both groups.



### Progress of reading comprehension

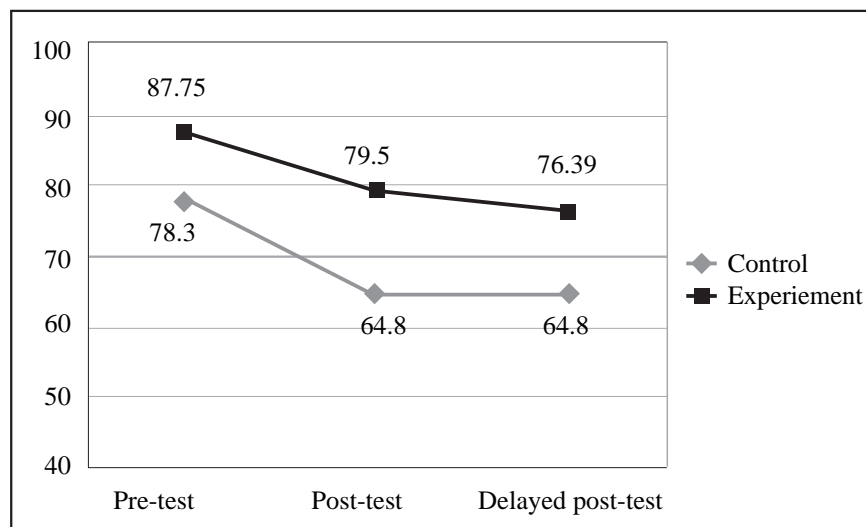
The three tests were divided into two categories—listening and reading—for verifying the progress of the two proficiencies during and after the research. At the reading part of the pretest, the mean scores were 78.30 and 87.75 (control and experimental groups, respectively) and the mean scores at the posttest were 64.80 and 79.50. The mean scores of the delayed posttest were 64.80 and 76.39, as shown in Table 3. Statistical significance ( $t = -2.553$ ,  $p = .016$  and  $t = -2.144$ ,  $p = .037$ ) was achieved when the reading scores were compared between the posttest and delayed posttest through the  $t$ -test. The mean reading score in the three tests is presented in Figure 8, and the reading comprehension in both groups exhibited a degradation during and after the research. The control group, Group 1, exhibited a major descent (from 78.30 to 64.80, 13.5 points less), and the performance remained the same after the experiment, with a score of 64.80. With the treatment of the video recommendation system, the reading comprehension in Group 2 degraded (from 87.75 to 79.50, 8.25 points less) during the experiment and the delayed posttest, as shown in Figure 8. The performance of the group also exhibited a decline from a score of 79.50 to 76.39 without the treatment of the video recommendation system.

**Table 3.** Descriptive data of  $t$ -test for reading. Group 1 is the control group, and Group 2 is the experimental group.

	Group	No	Mean	SD	$t$	Sig. (2-tailed)
Pre-test	1	20	78.30	24.583	-1.499	.144
	2	28	87.75	16.343		
Post-test	1	20	64.80	22.689	-2.553	.016*
	2	28	79.50	14.411		
Delayed Post-test	1	20	64.80	21.333	-2.144	.037*
	2	28	76.39	16.155		

\* $p < .05$ .

**Figure 8.** Line chart of three reading scores for both groups.

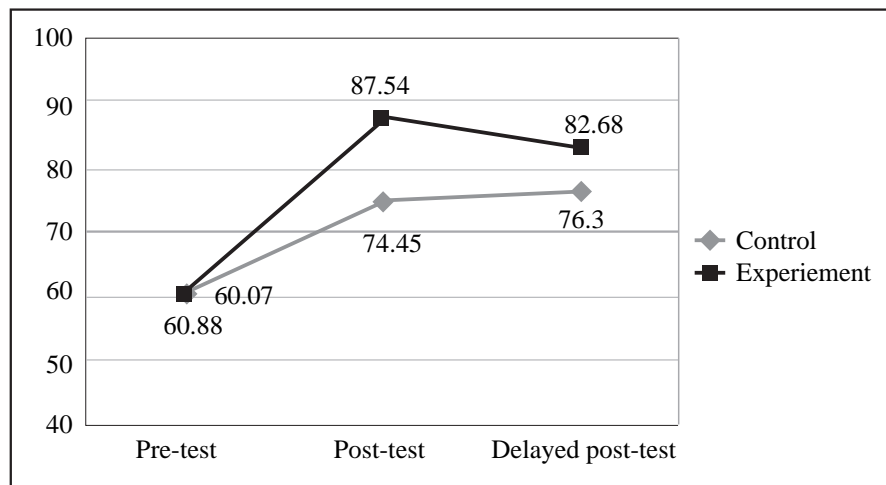


### Progress of listening proficiency

At the listening part of the pretest, the mean scores for both the groups were 60.88 and 60.07 (control and experiment), and there was no difference at the start of the research. The mean scores of listening at the posttest for both the groups were 74.45 and 87.54. The means of listening at delayed posttest were 76.30 and 82.68 (Table 4). Statistical significance ( $t = -2.870$ ,  $p = .006$ ,  $p < .01$ .) was noted in the scores between the two groups in the posttest. The listening score (Figure 9) during the experiment improved in both groups. The control group, Group 1, exhibited an increase in the score from 60.88 to 74.45 and then slight progress from 74.45 to 76.30 after the experiment. Overall, the control group exhibited an improvement (from 60.88 to 76.30) following the regular English instruction. In the experimental group, Group 2, with the treatment of the video recommendation system, the performance in listening proficiency exhibited an excessive advance (mean score from 60.07 to 87.54) during the experiment, but a degradation from a score of 87.54 to 82.68 without the treatment. Generally, the listening performance of the experimental group instructed with video recommendations based on learners' error responses greatly improved from a score of 60.07 to 82.68 and was higher than that of the control group.

**Table 4.** Descriptive data of *t*-test for listening performance. Group 1 is the control group, and Group 2 is the experimental group.

	Group	No	Mean	SD	<i>t</i>	Sig. (2-tailed)
Pre-test	1	20	60.88	16.721	.167	.868
	2	28	60.07	16.127		
Post-test	1	20	74.45	18.025	-2.870	.006**
	2	28	87.54	13.591		
Delayed Post-test	1	20	76.30	18.459	-1.117	.270
	2	28	82.68	20.195		

\*\*  $p < .01$ .**Figure 9.** Line chart of four listening scores for both groups.

### Analysis of learning achievement

Both the experimental and control groups were examined using the pretest and posttest. The scores in both tests were evaluated using the homogeneity of the variable test. The result of the homogeneity of the variable test indicated a  $p$ -value of .826 (Table 5) ( $F = .049$  and  $p = .826$ ). Because the  $p$ -value is greater than .05 homogeneities between the two groups cannot be rejected and we conclude that the variances were equal. Therefore, the scores of the experimental and control groups were homogenous. An analysis of covariates in ANCOVA ( $p = .000$ ,  $p < .001$ ) revealed

that the pretest and posttest scores for both groups were statistically significant. The control group did not show improvement (score from 139.18 to 139.50) during the experiment and then showed a slight improvement (score from 139.50 and 141.10) after the experiment. The experimental group exhibited excessive improvement (score from 147.83 to 167.04) with English instruction and the treatment of the video recommendation system based on learners' error responses.

**Table 5.** *Statistics of Homogeneity & ANCOVA between the groups*

	Sum of	df	Mean Square	<i>F</i>	Sig.
Homogeneity	15.698	1	15.698	.049	.826
ANCOVA	5017.523	1	5017.523	15.890	.000***

\*\*\*  $p < .001$

## Conclusion

The research results revealed that English instruction with video recommendations based on learners' error responses is more effective than regular English instruction. The results of this study specified the following. First, English instruction with video recommendations based on learner's error analysis is beneficial to English listening and reading proficiency. Compared with the case when regular English instructions were provided, learning was better in terms of listening and reading when the proposed method was used with the English instructions. Second, adaptive English learning based on a video recommendation system is more effective than regular reading and writing. Generally, the study results revealed that the group that was provided with English learning with the system exhibited higher performance. Moreover, it was found that learning English by using the video recommendation system based on learners' error responses helps to learn effectively. For the system that is developed using film scripts and error analysis, the instruction provided with the recommendation can be more widely applied to learners of all ages and proficiency levels. Through the audiovisual aid and the recommendation system, English instruction can be timeless and boundless to English learning. Providing English instructions using

the video recommendation system based on learners' error responses enhances learners' interests in practicing English, thus leading to learning effectiveness.

At the beginning of the study, students were informed that the English instructions contained various themes and issues given for assignments once in two weeks. This triggered their interest in learning English based on the study result and the instructional evaluation at the end of the semester. Moreover, the experiment was more satisfied with the instructions provided by the recommendation system to help them clarify their misunderstandings while listening or skimming through subtitles of videos. Moreover, based on the results of the study (scores from 147.83 to 167.04), the video recommendation system is an advantageous tool. The system improves the English listening ( $p = .006^{**}$ , *t-test* of posttest) and reading ( $p = .016^{*}$  and  $p = .037^{*}$ , *t-test* of posttest and delayed posttest) abilities of learners. Overall, the result is statistically significant ( $p = .000^{***}$ ).

Continuous progress in science and technology has provided learners with good resources to improve their skills. Adaptive language learning, such as learning using a recommendation system, has become a trend. In the past, the main recommendation methods were mostly based on keywords or terms and recommended to learners by searching for the similarity (Chen et al., 2018; Zeng et al., 2017). In line with many studies that utilize audiovisual aids to enhance learning and teaching (Chen, 2011; Hsu, 2018; Tschirner, 2001; Yang, 2011) the present study developed a recommendation system for supporting English instructions. In this study, English instructions were provided to learners, and they were asked to use a system that employed LSA to determine the potential implications of error responses given to questions with multiple choices. Based on the error responses, the video recommendation system provided video suggestions to learners. Through the recommendations based on error analysis, learners could solve the misunderstandings they had in English listening and reading.

Irrespective of whether the English instruction is provided with or without using the video recommendation system, the reading proficiency of both groups tended to decrease. The mean scores of reading of Groups 1 and 2 were 78.30 and 87.75 at the beginning of the study. The values drastically decreased to 64.80 and 79.50, respectively and the score of the experiment group decreased less than it of the control group

(8.25 less and 13.5 less, respectively). This study found that their reading proficiency deteriorated after they entered university. This might be attributed to the fact that the method of learning and teaching at a university is more self-directed and personalized, which is different from the rigorous and painstaking study in high school. Students at a university spend more time on student activities rather than studying all day long and how to trigger students' self-learning spontaneously is a crucial question in future research. Conversely, the videos with scripts were visual listening and reading materials (Hung & Chen, 2018; Tschirner, 2001) that reduced the anxieties of students while learning a language different from their mother tongue (Lee et al., 2017) and enhanced their interest in the target language patterns and authentication of the language use. The mean scores of listening in the pretest for both the groups were 60.88 and 60.07 at the beginning of the study and then drastically increased to 74.45 and 87.54, respectively.

While LSA is a well-equipped technology, the adaptive English learning system based on video recommendation is novel and the experimental results proposed with LSA provide reliable results. For future expanded research, the recommendation system might combine with another recommendation algorithm such as Probabilistic Latent Semantic Analysis (PLSA) or the neural network methods provided by Tensor Flow. Besides, the English instruction with video recommendation adopting a Lexile analyzer and the experts' suggestions to decide the video with proper oral languages, speaking pace, scenery, characters, and images, future research and further discussion will be the English instruction with video recommendation operated by Speech Recognition Techniques and Video Scene Recognition to scale the rating of videos.

## Acknowledgments

This paper has profited from all the students and the teachers who participated in the study and this manuscript was edited by Wallace Academic Editing. Without their assistance, the study would not have been completed. This research did not receive any specific grant from funding agencies in the public.



## References

- Altınay-Gazi, Z., & Altınay-Aksal, F. (2017). Technology as mediation tool for improving teaching profession in higher education practices. *Eurasia Journal of Mathematics, Science & Technology Education*, 13(3), 803-813. <https://doi.org/10.12973/eurasia.2017.00644a>
- Amiri, E., & Branch, L. (2012). A study of the application of digital technologies in teaching and learning English language and literature. *International Journal of Scientific & Technology Research*, 1(5), 103-107.
- Baker, F. B. (2001). *The basics of item response theory*. <http://ericae.net/irt/baker/software.htm>
- Bowyer - Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure?. *British Journal of Educational Psychology*, 75(2), 189-201. <https://doi.org/10.1348/000709904X22674>
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE-Life Sciences Education*, 6(1), 9-20. <https://doi.org/10.1187/cbe.06-12-0205>
- Cardoso, W. (2011). Learning a foreign language with a learner response system: The students' perspective. *Computer Assisted Language Learning*, 24(5), 393-417. <https://doi.org/10.1080/09588221.2011.567354>
- Chen, C. M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, 51(2), 787-814. <https://doi.org/10.1016/j.compedu.2007.08.004>
- Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51(2), 624-645. <https://doi.org/10.1016/j.compedu.2007.06.011>
- Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, 18(4), 341-364. <https://doi.org/10.1080/10494820802602329>
- Chen, C. M., Lee, H. M., & Chen, Y. H. (2005). Personalized e-learning system using item response theory. *Computers & Education*, 44(3), 237-255. <https://doi.org/10.1016/j.compedu.2004.01.006>
- Chen, H. J. H. (2011). Developing and evaluating SynctoLearn, a fully automatic video and transcript synchronization tool for EFL learners. *Computer Assisted Language*

- Learning*, 24(2), 117-130. <https://doi.org/10.1080/09588221.2010.526947>
- Chen, Y., Li, X., Liu, J., & Ying, Z. (2018). Recommendation system for adaptive learning. *Applied Psychological Measurement*, 42(1), 24-41. <https://doi.org/10.1177/0146621617697959>
- Conole, G. (2012). *Designing for learning in an open world* (Vol. 4). Springer Science & Business Media.
- Figueras, N., & Noijons, J. (Eds.) (2009). *Linking to the CEFR levels: Research perspectives*. [https://www.ealta.eu.org/documents/resources/Research\\_Colloquium\\_report.pdf](https://www.ealta.eu.org/documents/resources/Research_Colloquium_report.pdf)
- Forsyth, C. M., Luce, C., Zapata-Rivera, D., Jackson, G. T., Evanini, K., & So, Y. (2018). Evaluating English language learners' conversations: Man vs. machine. *Computer Assisted Language Learning*, 1-20. <https://doi.org/10.1080/09588221.2018.1517126>
- Futagi, Y., Deane, P., Chodorow, M., & Tetreault, J. (2008). A computational approach to detecting collocation errors in the writing of non-native speakers of English. *Computer Assisted Language Learning*, 21(4), 353-367. <https://doi.org/10.1080/09588220802343561>
- Gass, S. M. (2013). *Second language acquisition: An introductory course*. Routledge.
- Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education*, 63, 327-336. <https://doi.org/10.1016/j.compedu.2012.12.004>
- Hsu, M. H. (2008). A personalized English learning recommender system for ESL students. *Expert Systems with Applications*, 34(1), 683-688. <https://doi.org/10.1016/j.eswa.2006.10.004>
- Hsu, T. C. (2018). Using a concept mapping strategy to improve the motivation of EFL students in Google Hangouts Peer-Tutoring Sessions with native speakers. *Interactive Learning Environments*, 1-14. <https://doi.org/10.1080/10494820.2018.1463268>
- Hung, I. C., & Chen, N. S. (2018). Embodied interactive video lectures for improving learning comprehension and retention. *Computers & Education*, 117, 116-131. <https://doi.org/10.1016/j.compedu.2017.10.005>
- Kennedy, C., & Miceli, T. (2017). Cultivating effective corpus use by language learners. *Computer Assisted Language Learning*, 30(1-2), 91-114. <https://doi.org/10.1080/09588221.2016.1264427>
- Kuo, B. C., Liao, C. H., & Chang, C. J. (2018). Using LSA-based tools to enhance students' Chinese reading ability. *International Journal on Digital Learning Technology*, 10(1),

31-56.

- Lai, C., Shum, M., & Tian, Y. (2016). Enhancing learners' self-directed use of technology for language learning: The effectiveness of an online training platform. *Computer Assisted Language Learning*, 29(1), 40-60. <https://doi.org/10.1080/09588221.2014.889714>
- Lan, Y. J., Sung, Y. T., & Chang, K. E. (2007). A mobile-device-supported peer-assisted learning system for collaborative early EFL reading. *Language Learning & Technology*, 11, 130-151.
- Landauer, T. K., & Psotka, J. (2000). Simulating text understanding for educational applications with Latent Semantic Analysis: Introduction to LSA. *Interactive Learning Environments*, 8(2), 73-86. [https://doi.org/10.1076/1049-4820\(200008\)8:2;1-B;FT073](https://doi.org/10.1076/1049-4820(200008)8:2;1-B;FT073)
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3), 259-284. <https://doi.org/10.1080/01638539809545028>
- Laurillard, D. (2013). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.
- Lawley, J. (2016). Spelling: Computerised feedback for self-correction. *Computer Assisted Language Learning*, 29(5), 868-880. <https://doi.org/10.1080/09588221.2015.1069746>
- Lee, C. (2019). A study of adolescent English learners' cognitive engagement in writing while using an automated content feedback system. *Computer Assisted Language Learning*, 1-30. <https://doi.org/10.1080/09588221.2018.1544152>
- Lee, C., Cheung, W. K. W., Wong, K. C. K., & Lee, F. S. L. (2013). Immediate web-based essay critiquing system feedback and teacher follow-up feedback on young second language learners' writings: an experimental study in a Hong Kong secondary school. *Computer Assisted Language Learning*, 26(1), 39-60. <https://doi.org/10.1080/09588221.2011.630672>
- Lee, C., Wong, K. C., Cheung, W. K., & Lee, F. S. (2009). Web-based essay critiquing system and EFL students' writing: A quantitative and qualitative investigation. *Computer Assisted Language Learning*, 22(1), 57-72. <https://doi.org/10.1080/09588220802613807>
- Lee, C., Yeung, A. S., & Ip, T. (2017). University English language learners' readiness to use computer technology for self-directed learning. *System*, 67, 99-110. <https://doi.org/10.1016/j.system.2017.05.001>
- Lennon, C., & Burdick, H. (2004). *The Lexile framework as an approach for reading measurement and success*. <http://www.lexile.com>

- Li, J. F., & Huang, C. T. (2008). The correlation between EFL learners' vocabulary size and reading performance in a communication-oriented language proficiency test. *SPECTRUM: NCUE Studies in Language, Literature, Translation*, 3, 131-148. <https://doi.org/10.29601/YYWYLL.200807.0010>
- Lincoln, T. D. (2011). Students' experiences of e-learning in higher education: The ecology of sustainable innovation-by Robert A. Ellis and Peter Goodyear. *Teaching Theology & Religion*, 14(1), 79-80. <https://doi.org/10.1111/j.1467-9647.2010.00678.x>
- Luckin, R. (2010). *Re-designing learning contexts: Technology-rich, learner-centred ecologies*. Routledge.
- Martin, D. I., & Berry, M. W. (2007). Mathematical foundations behind latent semantic analysis. *Handbook of Latent Semantic Analysis*, 35-56. <https://doi.org/10.4324/9780203936399.ch2>
- Neville, D. O., Shelton, B. E., & McInnis, B. (2009). Cybertext redux: Using digital game-based learning to teach L2 vocabulary, reading, and culture. *Computer Assisted Language Learning*, 22(5), 409-424. <https://doi.org/10.1080/09588220903345168>
- North, B. (2007). The CEFR illustrative descriptor scales. *The Modern Language Journal*, 91(4), 656-659. [https://doi.org/10.1111/j.1540-4781.2007.00627\\_3.x](https://doi.org/10.1111/j.1540-4781.2007.00627_3.x)
- Roschelle, J., Dimitriadis, Y., & Hoppe, U. (2013). Classroom orchestration: Synthesis. *Computers & Education*, 69, 523-526. <https://doi.org/10.1016/j.compedu.2013.04.010>
- Shen, M. Y., & Wu, W. S. (2009). Technical university EFL learners' reading proficiency and their lexical inference performance. *Electronic Journal of Foreign Language Teaching*, 6(2), 189-200.
- Shu, J., Shen, X., Liu, H., Yi, B., & Zhang, Z. (2018). A content-based recommendation algorithm for learning resources. *Multimedia Systems*, 24(2), 163-173. <https://doi.org/10.1007/s00530-017-0539-8>
- Smith, M., & Turner, J. (2016). The Common European Framework of Reference for languages (CEFR) and the Lexile® framework for reading. *Meta Metrics*. [https://metametricsinc.com/wpcontent/uploads/2018/01/CEFR\\_1.pdf](https://metametricsinc.com/wpcontent/uploads/2018/01/CEFR_1.pdf)
- Smith, M., Turner, J., Sanford-Moore, E., & Koons, H. H. (2016). The Lexile framework for reading: An introduction to what it is and how to use it. In *Pacific Rim Objective Measurement Symposium (PROMS) 2015 Conference Proceedings* (pp. 411-424). Springer Singapore.
- Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1965). *The grammatical structures of English and Spanish* (Vol. 4). University of Chicago Press.
- Sun, S. Y. (2018). Student configuration and place-making in fully online language

- learning. *Computer Assisted Language Learning*, 31(8), 1-28. <https://doi.org/10.1080/09588221.2018.1466808>
- Sung, Y. T., Liao, C. N., Chang, T. H., Chen, C. L., & Chang, K. E. (2016). The effect of online summary assessment and feedback system on the summary writing on 6th graders: The LSA-based technique. *Computers & Education*, 95, 1-18. <https://doi.org/10.1016/j.compedu.2015.12.003>
- Tinkham, T. (1993). The effect of semantic clustering on the learning of second language vocabulary. *System*, 21(3), 371-380. [https://doi.org/10.1016/0346-251X\(93\)90027-E](https://doi.org/10.1016/0346-251X(93)90027-E)
- Tschirner, E. (2001). Language acquisition in the classroom: The role of digital video. *Computer Assisted Language Learning*, 14(3-4), 305-319. <https://doi.org/10.1076/call.14.3.305.5796>
- Velleman, E., & van der Geest, T. (2014). Online test tool to determine the CEFR reading comprehension level of text. *Procedia Computer Science*, 27, 350-358. <https://doi.org/10.1016/j.procs.2014.02.039>
- Wang, Y. H., & Liao, H. C. (2011). Data mining for adaptive learning in a TESL-based e-learning system. *Expert Systems with Applications*, 38(6), 6480-6485. <https://doi.org/10.1016/j.eswa.2010.11.098>
- Yang, Y. F. (2011). Engaging students in an online situated language learning environment. *Computer Assisted Language Learning*, 24(2), 181-198. <https://doi.org/10.1080/09588221.2010.538700>
- Zeng, J., Ge, J., Zhou, Y., Feng, Y., Li, C., Li, Z., & Luo, B. (2017, November). Statutes recommendation based on text similarity. In *Web Information Systems and Applications Conference (WISA), 2017 14th* (pp. 201-204). IEEE. <https://doi.org/10.1109/WISA.2017.52>

