Chen, C.-H., Koong, C.-S., & Liao, C. (2022). Influences of Integrating Dynamic Assessment into a Speech Recognition Learning Design to Support Students' English Speaking Skills, Learning Anxiety and Cognitive Load. *Educational Technology & Society*, *25* (1), 1-14.

Influences of Integrating Dynamic Assessment into a Speech Recognition Learning Design to Support Students' English Speaking Skills, Learning Anxiety and Cognitive Load

Chih-Hung Chen[1*], Chorng-Shiuh Koong[2] and Chien Liao[1]

[1]Master Program of Professional Teacher, National Taichung University of Education, Taiwan // [2]Department of Computer Science, National Taichung University of Education, Taiwan // duke.chchen@gmail.com // csko@mail.ntcu.edu.tw // btp107203@gm.ntcu.edu.tw

[*]Corresponding author

(Submitted January 6, 2021; Revised February 4, 2021; Accepted February 23, 2021)

ABSTRACT: Artificial intelligence (AI) technology has been progressively utilized in educational environments in recent years, due to the advances in computing and information processing techniques. The automatic speech recognition technique (ASR) provides students with instantaneous feedback and interactive oral practice for supporting a context with self-paced learning. Corrective feedback (CF) should be combined with ASR-based systems to enhance students' speaking performance, and to reduce their cognitive load. However, learners' perceptions of CF are mixed, and CF might give rise to learning anxiety. In this study, a dynamic assessment-based speech recognition (called DA-SR) learning system was designed to facilitate students' English speaking. Moreover, a quasi-experiment was implemented to evaluate the effects of the proposed approach on students' speaking learning effectiveness, via respectively providing the DA-SR and the corrective feedback-based speech recognition (called CF-SR) approaches for the experimental and control groups. The experimental results revealed that both the DA-SR group and the CF-SR group can effectively improve the students' English speaking skills, and decrease their English speaking learning anxiety. Moreover, this study further demonstrated that the DA-SR approach successfully reduced students' English class performance anxiety, and extraneous cognitive load in comparison with the CF-SR approach. It could be a valuable reference for designing English speaking learning activities in EFL learning environments.

Keywords: Artificial intelligence, Speech recognition, Corrective feedback, Dynamic assessment, Learning anxiety

1. Introduction

English is regarded as the most widely spoken language in the world. With globalization and the rapid advances in technology, English is now in widespread use, highlighting the importance of enhancing students' English competencies and global perspectives (Chen, 2020; Fu et al., 2019). Foreign language learning can be probed according to the four language skills, namely listening, reading, writing and speaking. Among these skills, enhancing speaking ability is widely deemed to be a difficult task for most English as a foreign language (EFL) learners (Gan, 2013). A great number of studies have investigated instructional approaches or contexts of EFL to promote students' English speaking skills and learning motivations (Abdullah et al., 2019). For example, Chien et al. (2020) adopted the peer assessment strategy in a spherical video-based virtual reality environment for situating learners in an authentic learning situation, and for directing them to comment on peers' English speaking performance. Such a learning strategy effectively enhanced the learners' English speaking skills and facilitated their reflections on what they had learned. Furthermore, a learner may feel anxious about public speaking or about answering questions (Bodnar et al., 2017), and accordingly reduction in learning anxiety has been regarded as a crucial factor for improving students' English speaking skills (Liu & Jackson, 2008; Zhang & Liu, 2018). Chen and Hwang (2020) asserted that speaking anxiety is related to language development, and anxiety may affect learners' oral competence. Thus, the provisions of feedback guidance and reductions in learning anxiety have been considered as significant factors for improving students' English speaking skills.

With the rapid development of information and communication technology, the ways to learn languages have changed. Language learning materials can be displayed in an interactive manner with multimedia (Hwang & Fu, 2019). Over the last few decades, computer-assisted language learning (CALL) involving diverse computer-mediated activities has attracted much attention (Fathi & Ebadi, 2020). Taking advantage of technology in an EFL class is able to facilitate practical language skills, and to reduce learning anxiety about speaking mistakes via individual practice (de Vries et al., 2015; Kuru Gönen, 2019). It is suggested that new instructional strategies or tools should be adopted to support EFL learners in promoting language skills and encouraging more interaction (Yang & Kuo, 2020). Moreover, artificial intelligence (AI) technology has been progressively utilized in educational environments in recent years, due to the advances in computing and information processing techniques. AI aims to deal with cognitive problems which are related to human intelligence, and specifically Artificial Intelligence in Education (AIED) refers to the application of AI technologies in educational contexts to facilitate instruction or decision making, such as intelligent tutoring systems and adaptive learning systems (Chen et al., 2020a; Hwang et al., 2020b). AIED can be defined from both broad and narrow perspectives, namely the use of AI techniques in education, and the utilization of machine learning (ML) or deep learning (DL) techniques in education, respectively (Chen et al., 2020b).

AIED has offered new opportunities for facilitating superior technology-enhanced learning contexts and for carrying out productive learning activities, for example, the provision of personalized learning guidance or the supply of individual needs (Chen et al., 2020a; Hwang et al., 2020b). The automatic speech recognition technique (ASR), powered by DL neural networking, refers to a kind of technology which synchronously transcribes text streams from individual speech (Shadiev et al., 2018). With the popularity of mobile devices, the adoption of ASR in EFL speaking courses is gradually increasing (Nguyen et al., 2018). ASR can provide students with instantaneous feedback and interactive oral practice for supporting a context with self-paced learning (Luo, 2016). However, ASR faces the same issue regarding cognitive overload as spoken production does. It has been asserted that CALL systems or ASR-based systems should offer automatic corrective feedback (CF) for enhancing students' speaking performance, and for reducing their cognitive load (de Vries et al., 2015; Young & Wang, 2014). Moreover, some ASR drawbacks regarding hardware and software have been described (Crescenzi-Lanna, 2020; Yang & Meinel, 2014), indicating the necessity of investigating learners' perceptions of the utilization of an ASR-based learning system.

CF in second language acquisition refers to the responses to the correctness or appropriateness of a learner's production or comprehension (Li, 2010; Li & Vuono, 2019), which is capable of providing students with both the opportunity and time for self-repaired output (de Vries et al., 2015). CF has played a vital role in the type of scaffolding that teachers offer for improving students' EFL learning (Lyster et al., 2013). Some previous research has illustrated the influences of oral CF on students' speaking skills in EFL learning, and has proposed several types of feedback to enhance students' speaking abilities (Couper, 2019). For example, Lyster and Ranta (1997) illustrated a taxonomy of six types of different corrective feedback that teachers could adopt in the classroom, namely explicit correction, recasts, clarification requests, metalinguistic feedback, elicitation, and repetition. These corrective types can be categorized into implicit and explicit feedback. Moreover, some internal and external learner variables have been proven to be significant in determining the effectiveness of CF (Penning de Vries et al., 2020). Individual learners' proficiency, motivation, and anxiety are considered internal variables, whereas learning contexts are deemed as external variables, such as CF type, outcome measures, and CALL. It is asserted that learners' perceptions of CF are mixed, and CF might give rise to learning anxiety (Bodnar et al., 2017). This implies the value of designing technology-enhanced speaking instruction, and of providing proper CF in the EFL classroom (Rassaei, 2019). Furthermore, working memory capacity is also considered to be crucial to the effects of CF, indicating the importance of probing learners' cognitive load during English speaking activities (Penning de Vries et al., 2020).

Dynamic assessment (DA), which is a kind of alternative assessment, has been referred to as a useful interactive pedagogical approach (Cho et al., 2020). One key component of DA is scaffolded feedback, which is displayed in some form of corrective feedback (Herazo et al., 2019). DA depicts learners' cognitive structures so as to diagnose learner difficulties, and to recognize potential improvements (Allal & Ducrey, 2000; Wang & Chen, 2016). For example, Antón (2009) utilized the DA approach to evaluate students' actual and emergent abilities, and facilitated the programming of individualized instruction. Furthermore, Rezaee et al. (2020) explored the potential effects of a mobile-based dynamic assessment on EFL learners' oral fluency, and verified that the students' speaking fluency was enhanced by the proposed approaches with both text-chat, and voice-chat contexts.

Collectively, in this study, a speech recognition approach with dynamic assessment was proposed. Based on the approach, a dynamic assessment-based speech recognition (called DA-SR) system was designed to facilitate students' English speaking. Furthermore, a quasi-experiment was implemented to evaluate the effects of the proposed approach on students' speaking learning effectiveness, via respectively providing the DA-SR and the corrective feedback-based speech recognition (called CF-SR) for the experimental and control groups. The research questions in this study are listed below.

- Do the students who learned with the DA-SR approach outperform those who learned with the CF-SR approach in terms of their English speaking skills?
- Do the students who learned with the DA-SR approach reveal a lower degree of learning anxiety than those who learned with the CF-SR approach?
- Can the DA-SR approach reduce the students' cognitive load in comparison with the CF-SR approach?

2. Literature review

2.1. Automatic speech recognition for English speaking skills

The automatic speech recognition technique (ASR) is deemed as a potentially valuable AI technology which can facilitate intelligible English speech of EFL students by means of immediately transcribing text streams from their speech (Huang et al., 2016; Shadiev et al., 2014). Several previous studies have emphasized the potentiality of integrating ASR with CALL for pronunciation learning (Young & Wang, 2014), such as reduced anxiety for non-native speaking (de Vries et al., 2015), positive learning motivation (Nguyen et al., 2018), and their speaking skills in the foreign language (Wang & Young, 2014). For example, Cavus and Ibrahim (2017) adopted a speech recognition technology on mobile devices for recognizing and correcting students' spoken words, and the research results revealed that the developed learning system significantly enhanced the students' English learning skills.

Information technology offers the function of repeated training, and expands the opportunities for utilizing the target language. Moreover, an individual can repeatedly conduct English speaking practices using the ASR technology, so as to improve their fluency in English (Wang & Young, 2014). ASR-based learning systems can provide students with opportunities and integrated learning stimulation for promoting their non-native oral skills via immediately evaluating English utterances (Chen, 2011). Furthermore, the integration of the ASR-based learning system affords individualized and instantaneous feedback for creating a learning context in which individual students can learn at their own pace (Luo, 2016). Hsu (2016) described that an ASR-based learning system is able to facilitate students' metacognitive strategies in language learning via offering them timely feedback, resulting in the enhancement of their pronunciation.

With the advances in mobile and wireless technology, mobile devices have great potentiality for pedagogical application in language learning (Zhang & Zou, 2020). Via the advancement of the mobile-assisted language learning systems, the significant advantages of ASR in improving EFL learners' speaking proficiency have drawn much attention (Ahn & Lee, 2016). Such a learning context is capable of reducing students' speaking anxiety for foreign English by way of providing repeated drills and self-paced learning at any time, leading to an unpressured speaking environment (Wang & Young, 2014). Moreover, students may feel anxious about speaking out in front of classmates in class. Situating them in an unpressured speaking environment using the ASR technology is capable of reducing anxiety for foreign English speaking, indicating that ASR-based CALL systems have the potential to implement excellent English speaking and conversation practice situations (de Vries et al., 2015).

Some previous research has illustrated that spoken production requires control of the articulatory system, and may lead to great cognitive load (de Vries et al., 2015). Cognitive load refers to a multidimensional construct of the cognitive system regarding the load while performing a particular task (Paas et al., 2003; Paas & van Merriënboer, 1994). Intrinsic cognitive load is considered as an inherent component of the materials itself and individual degree of prior experience, while extraneous cognitive load originates in the excess information processing caused by the instructional design (Leahy & Sweller, 2016; Wu et al., 2018). Due to the restricted working memory capacity of learners, it is crucial to explore the relation between an instructional design and cognitive load, so as to accommodate the difficulty level of the learning activities to students' learning capabilities (Hwang et al., 2020a; Lai et al., 2019). Several previous studies have asserted the significance of providing learners with automatic corrective feedback in dominating cognitive load, while adopting a CALL system (de Vries et al., 2015). Although related studies have revealed that students produced more accurate utterances with the support of corrective feedback, few have evaluated the feedback design of the ASR-based learning systems, due to lacking sufficient pedagogical approaches for the feedback provision (Young & Wang, 2014). Thus, it is critical to carefully design the presentation of corrective feedback in an ASR-based system for promoting students' speech skills, and for reducing their cognitive load.

2.2. Dynamic assessment

Dynamic assessment (DA) is one kind of alternative assessment which integrates teaching and assessment into an interactive pedagogical approach with the provision of suitable forms of mediation (Cho et al., 2020; Ebadi & Rahimi, 2019). DA aims to portray a more complete image of learners' cognitive structures for enhancing the diagnosis of students' learning difficulties and for recognizing the developmental trajectory, by means of directly measuring their replies to specific interventions (Allal & Ducrey, 2000; Wang & Chen, 2016). DA is capable of promoting learners' achievements and of probing potential abilities via offering the details of their abilities to develop the intervention programs (Swanson & Lussier, 2001). For example, Antón (2009) declared that DA empowers a deeper characterization of learners' actual and latent abilities, and advances individualized instruction that can adapt to individual needs.

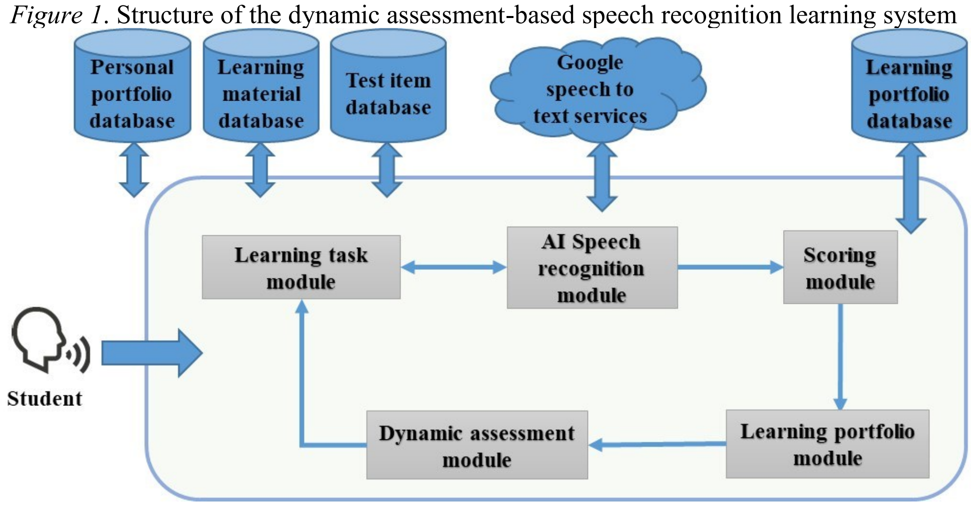
Previous research has illustrated the potential benefits of DA for improving students' learning effectiveness. Several related studies have probed the advantages of DA from the perspective of Vygotsky's zone of proximal development (ZPD), concentrating on developable abilities via intervention and interaction (Antón, 2009; Ebadi & Rahimi, 2019). For example, Wu et al. (2017) revealed that computerized dynamic adaptive tests are an effective approach for promoting learning achievement by providing individualized prompts. Bakhoda and Shabani (2019) designed a program with three sets of visual/audio/textual prompts (implicit to explicit) for evaluating emerging ZPD, and concluded that adapting to personal learning preferences with fine-tuned mediations in a computerized DA is practical. Furthermore, Rezaee et al. (2019) explored the impact of a mobile-based dynamic assessment on EFL students' oral accuracy, and declared that the proposed approach significantly improved students' oral accuracy. Andujar (2020) illustrated that DA and the dialogic mediation facilitated students' reflection on language performance, resulting in less requirement for explicit feedback and explanations.

Considering all of this evidence, it was revealed that ASR has been considered as an effective approach for enhancing students' EFL learning. However, on account of insufficient pedagogical methods for the feedback provision, few related studies have probed into the feedback mechanism designed for the ASR-based learning systems. It still remains a crucial issue to investigate the effects of integrating DA into an ASR learning context. Accordingly, this study aimed to develop a dynamic assessment-based speech recognition approach and to utilize it in an elementary school English course to evaluate its effects on students' English speaking skills, learning anxiety and cognitive load.

3. Development of the dynamic assessment-based speech recognition learning system

In this study, a dynamic assessment-based speech recognition (DA-SR) system was designed via integrating the dynamic assessment mechanism into speech recognition for enhancing students' English speech in an elementary school English course. The system was implemented utilizing PHP, MySQL, HTML, JavaScript and Google speech to text. Moreover, each student was furnished with a tablet computer and a headset for learning with no limits of time or space. Figure 1 depicts the structure of the DA-SR learning system which consists of the learning task module, the AI speech recognition module, the scoring module, the learning portfolio module, and the dynamic assessment module. Furthermore, some databases are set up to assist the modules, such as the learning material, the personal profile, the task item and the learning portfolio databases.

Figure 1. Structure of the dynamic assessment-based speech recognition learning system



The interface of the DA-SR learning system is depicted in Figure 2, which consisted of the number of the task, timing, illustration of the task, speech recognition, the prompt, and the submit button. Three kinds of learning tasks were designed in the learning activity, that is, picture reading, sentence pattern reading, and short conversations. The DA-SR learning procedure is portrayed in Figure 3. After an individual logs into the learning system, the learning task is displayed. For example, a picture of a zebra and a question, "What do you see?" are shown in the "short conversations" task. When an individual presses the "start" button and says an answer, the

DA-SR learning system immediately displays the text from the transcription of spoken language by speech recognition and requests the individual to confirm the transcription (as shown in Figure 4).

Figure 2. The interface of the DA-SR learning system

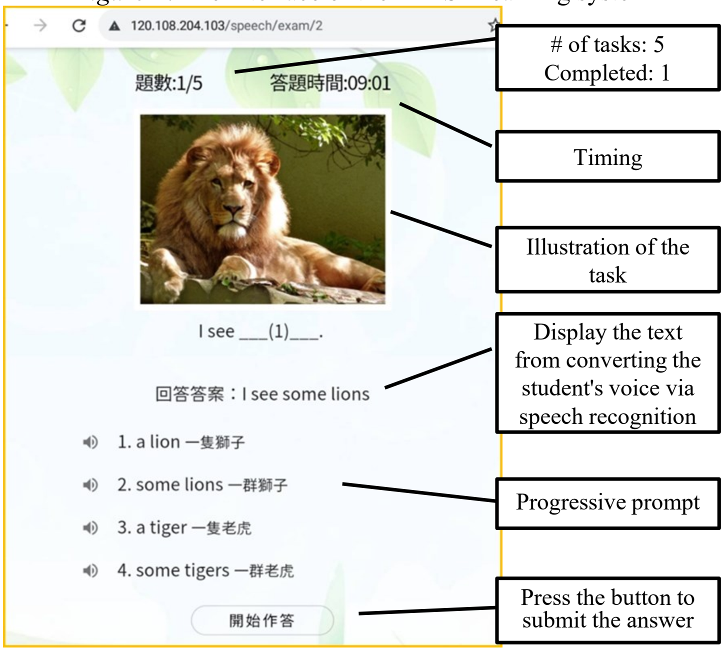
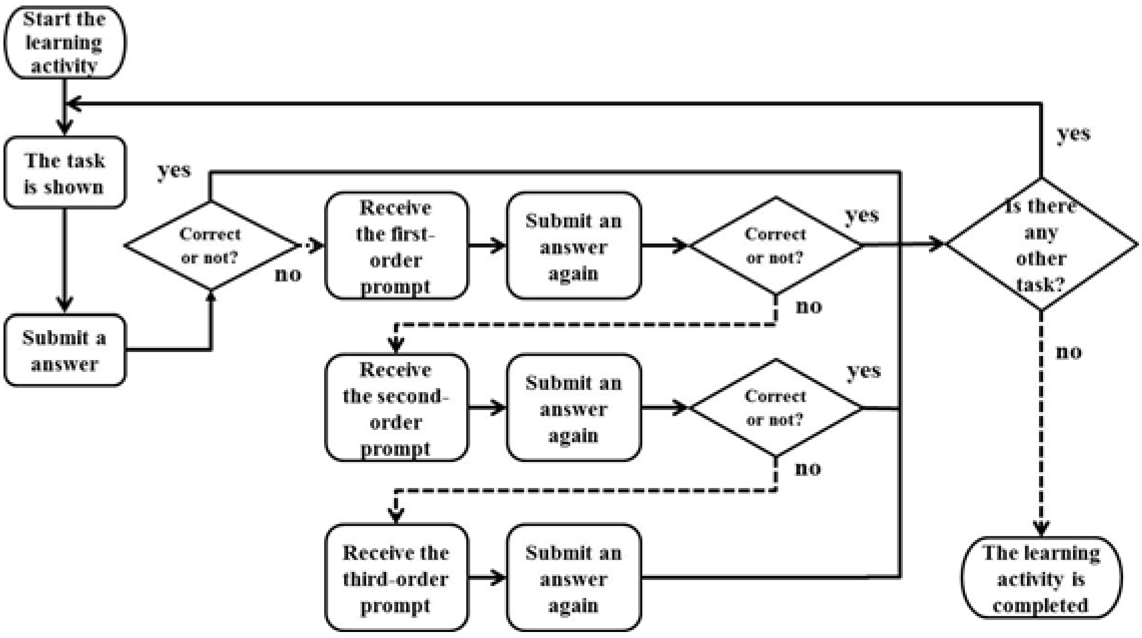


Figure 3. The learning process of the progressive prompting dynamic assessment



If the participant fails to give an appropriate answer, the DA-SR learning system assists her/him in accomplishing the task by utilizing a dynamic assessment approach. The more times a participant fails, the more concrete prompts that are given to her/him. As depicted in Figure 5, when a participant fails to produce a proper sentence the first time, the learning system provides four prompt items related to the grammar or dialogue context of the appropriate answers as the first-order prompt. If the participant fails to produce a proper sentence again, the learning system offers the Chinese translation and the application context of the four items as the second-order prompt. Furthermore, if the participant still could not submit a fitting answer the third time, the learning system then provides an audio file for demonstrating a suitable sentence.

Regarding the "picture reading" task, a picture (e.g., monkeys) is displayed on the mobile device, and the participant needs to say an answer. The three-order prompts are the provision of four prompt items (i.e., monks / monsters / monkeys / money) similar to the pronunciation of the correct answer, the supply of the Chinese translation of the prompt items, and the support of four audio files of the items in sequence. Furthermore, with respect to the "sentence pattern reading" task, a picture (e.g., a lion) and an incomplete sentence (e.g., "I see ___") are revealed on the screen. The three-order prompts are the aid of four prompt items close to the correct word or phrase (i.e., a lion / some lions / a tiger / some tigers), the support of the Chinese translation of the prompt items, and the assistance of providing four related audio files in turn. The more prompts the participants need to produce a proper sentence, the lower score they will receive. Upon successfully completing a learning task, the participant can move to the next task. When all the learning tasks are accomplished, the learning activity is completed.

Figure 4. Illustration of the transcription of spoken language

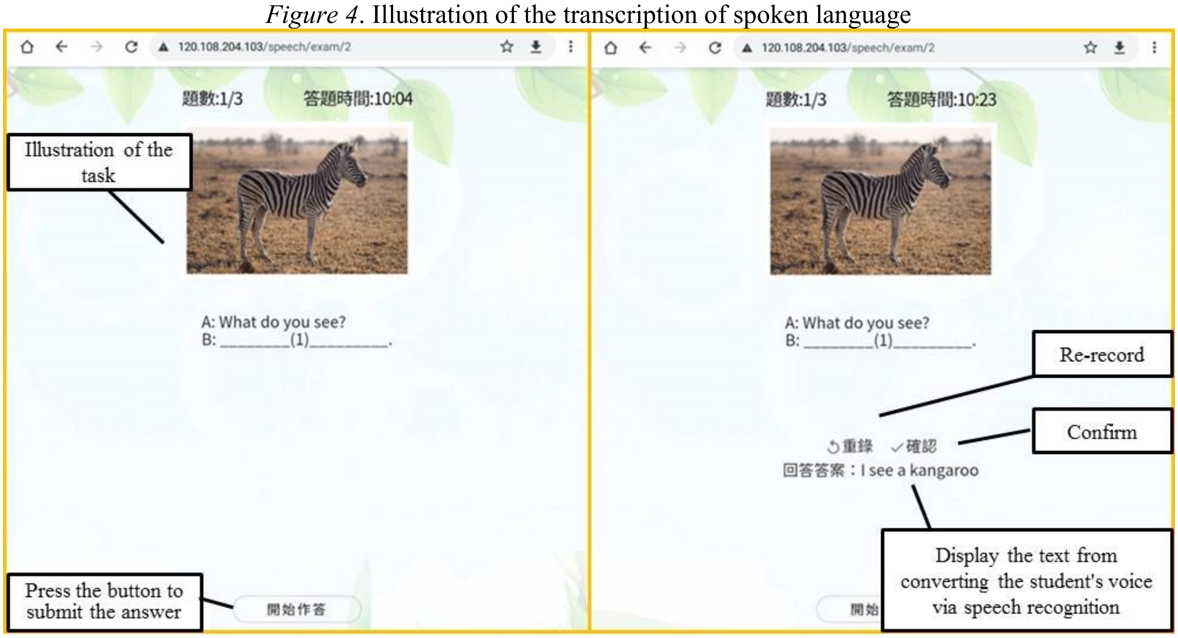
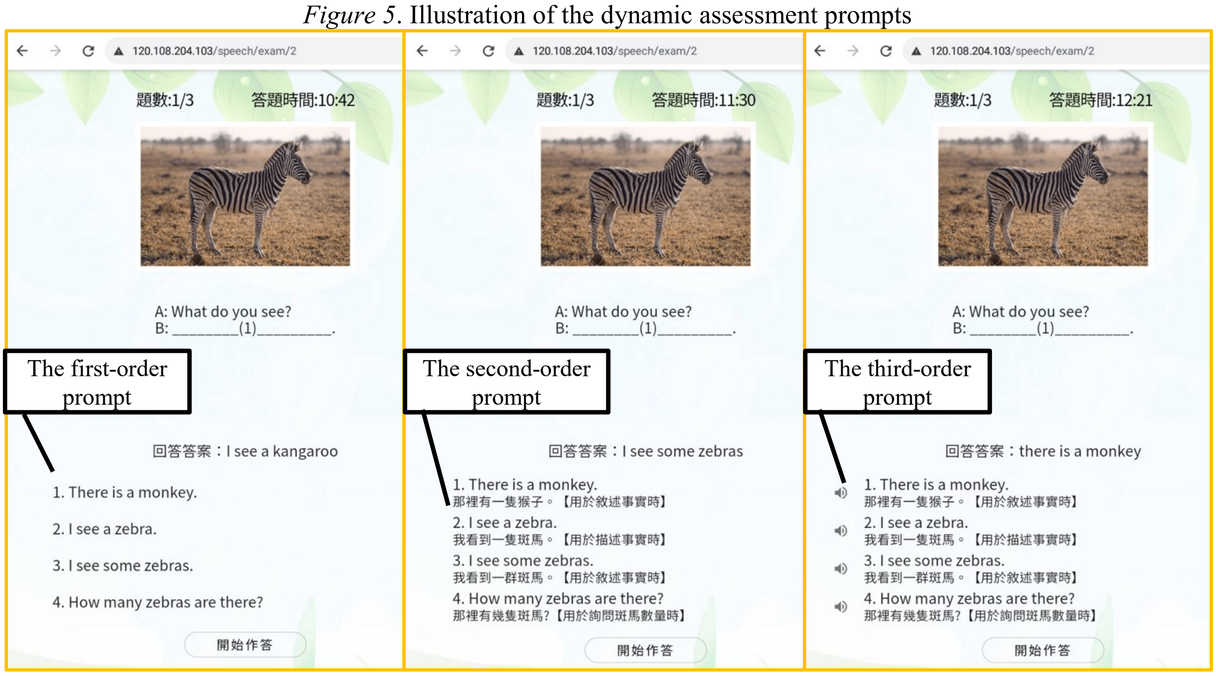


Figure 5. Illustration of the dynamic assessment prompts



4. Method

4.1. Participants

A total of 56 students from four classes of fifth graders (10- or 11-year-olds) in an elementary school in middle Taiwan were recruited for the experiment. They had English classes for three periods (a period of 40 minutes) per week. Among the four classes, two were allocated to be the experimental group (called the DA-SR group, *n* = 30), learning English speaking skills with the dynamic assessment-based speech recognition; the other two were the control group (called the CF-SR group, *n* = 26), learning English speaking skills with the corrective feedback-based speech recognition. All students in this study were already familiar with mobile technology-assisted learning.

4.2. Experimental procedure

In this study, the different English learning activities were designed to investigate the influences of integrating the dynamic assessment into a speech recognition design on the students' English speaking skills and perceptions. The experimental activity was conducted in a regular elementary school English class, and the experimental procedure is portrayed in Figure 6. First of all, the two groups experienced a regular 4-week English unit, and took the pre-test regarding English speaking skills, and filled out the pre-questionnaire about their learning anxiety.

Figure 6. The experimental procedure

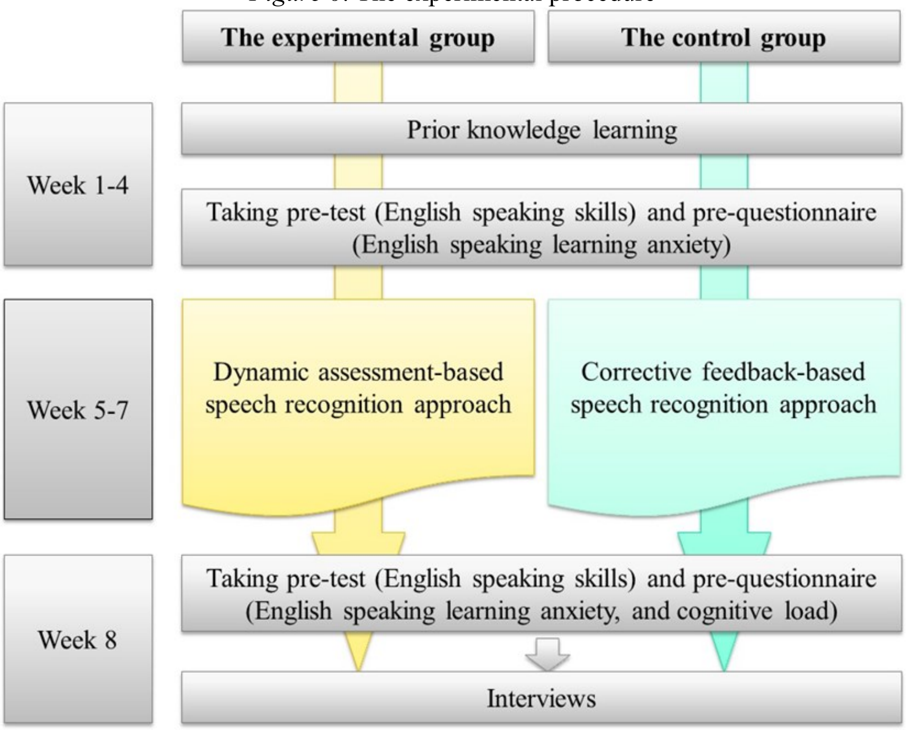


Figure 7. Snapshots of the DA-SR learning activity

Afterwards, the two groups conducted the English speaking activities with different learning approaches over 3 weeks. The learning materials, the learning tasks and the prompting content offered for the two groups were all identical, whereas the ways to display the prompts were different. When a wrong answer was selected, the learning systems activated the prompting functions. The DA-SR group was guided with the prompting content in three stages, while the CF-SR group was provided with the full prompting content all at once. Figure 7 depicts the snapshots of the learning scenarios regarding the DA-SR learning activity.

Upon completing the learning tasks, all students took the post-test concerning English speaking skills, and filled out the post-questionnaires of learning anxiety and cognitive load. Finally, a one-on-one interview was executed to investigate the views of six students recruited from each group.

4.3. Measuring tools

In this study, the pre- and post-tests of English speaking skills, and the questionnaires of learning anxiety and cognitive load were used as the instruments for assessing the students' English learning.

The pre- and post-tests were developed to evaluate the students' English speaking skills. Both tests comprised three kinds of questions with 16 items, namely short-answer questions, fill-in-the-blank items and short-answer questions about a photograph. Example items for the three kinds of questions are: "How many members are there in your family?" "_____ _____ _____ _____ in the sky" and "Are you going to the museum?" Both tests were scored on a scale of 0-80, based on the "Pre A1 Starters" assessment scale that is the first of three Cambridge English Qualifications. The assessment is formed with three criteria which are defined in candidate behaviour. The two experts who had more than 5 years' experience of teaching English courses designed the pre- and post-tests.

The learning anxiety questionnaire was modified from the instrument in Thompson and Lee's (2013) study. The original instrument consisted of four dimensions, namely "English class performance anxiety," "lack of self-confidence in English," "confidence with native speakers of English," and "fear of ambiguity in English." Furthermore, due to the learning context, the "confidence with native speakers of English" dimension and some items in the other dimensions were excluded in this study. Eventually, the adapted questionnaire of learning anxiety was composed of three scales with 18 items. The "English class performance anxiety" scale was made up of eight items (e.g., "In English class, I am so nervous that I forget what I know"). The "lack of self-confidence in English" scale included three items (e.g., "I keep considering that the other classmates speak English better than I do"). The "Fear of ambiguity in English" scale comprised seven items (e.g., "I always feel anxious about English class, although I am well prepared for it."). All items utilized a 5-point Likert rating scheme, and reverse scoring was used to re-code the responses for transforming a low point into the corresponding high point on the questionnaire. The higher the score the participants chose, the higher English learning anxiety they felt. The Cronbach's α values of the three dimensions computed by the adapted version were 0.88, 0.89, and 0.89, respectively, presenting highly acceptable reliability for rating students' English learning anxiety.

The cognitive load questionnaire was adopted from the instrument developed by Hwang et al. (2013). It had two dimensions using a 5-point Likert scale, including "mental load" and "mental effort." The mental load dimension comprised five items, while the mental effort dimension included three items. Two example items respectively for the "mental load" and "mental effort" dimensions are: "It was difficult for me to comprehend the learning content in the activity" and "It was difficult for me to follow and realize the instructional approach in the learning activity." The Cronbach's alpha coefficients of the two dimensions described by the original study were 0.85 and 0.86, respectively, showing highly acceptable reliability in internal consistency.

5. Experimental results

5.1. English speaking skills

One of the objectives of this study was to compare the impact of the DA-SR approach and that of the CF-SR approach on students' English speaking skills. Firstly, Spearman's rank correlation coefficients were computed between two sets of two experts who were recruited to judge the students' English speaking skills based on the "Pre A1 Starters" assessment scale. The Spearman's rho coefficients of the pre- and post-tests were 0.94 ($p < .01$) and 0.89 ($p < .01$), respectively, showing excellent intercoder reliability.

The paired t tests were executed to individually investigate the effects of the two learning approaches on the students' English speaking skills. Regarding the students' skills with the DA-SR approach, a significant difference was confirmed between the two tests with $t = -2.77$ ($p < .01$), as shown in Table 1. The means of the students' English speaking skills for the pre- and the post-tests respectively were 50.87 ($SD = 17.46$) and 57.50 ($SD = 17.57$). It was verified that the students who adopted the DA-SR approach significantly promoted their English speaking skills. On the other hand, with respect to the CF-SR approach, a significant difference was found between the two tests with $t = -2.71$ ($p < .05$), as depicted in Table 2. The means of the students' English speaking skills for the pre- and post-tests respectively were 41.35 ($SD = 24.24$) and 48.54 ($SD = 20.68$). It was evidenced that the students who learned with the CF-SR approach significantly improved their English speaking skills. Accordingly, both the DA-SR and the CF-SR approaches were of benefit to the students' English speaking skills.

Table 1. The paired *t*-test result of the experimental group's English speaking skills

Variable and source	*n*	Mean	*SD*	*t*
Pre-test skill	30	50.87	17.46	-2.77^{**}
Post-test skill	30	57.50	17.57	

Note. $^{**}p < .01$; Both tests were scored on a scale of 0-80.

Table 2. The paired *t*-test result of the control group's English speaking skills

Variable and source	*n*	Mean	*SD*	*t*
Pre-test skill	26	41.35	24.24	-2.71^{*}
Post-test skill	26	48.54	20.68	

Note. $^{*}p < .05$; Both tests were scored on a scale of 0-80.

Furthermore, a one-way ANCOVA was adopted to probe the influence of the different learning approaches on students' English speaking skills by excluding the interference from the two groups' prior skills. The pre-test skills were used as a covariate, while the learning approach and the post-test skills were respectively utilized as an independent variable and a dependent variable. Firstly, the homogeneity test was executed to evaluate the appropriateness of the utilization of the ANCOVA. It was proven that the assumption of homogeneity of regression was not violated with ($F = 0.01$, $p > .05$), and subsequently the ANCOVA was conducted. As shown in Table 3, no significant difference was found between the two groups' English speaking skills ($F = 0.43$, $p > .05$, $\eta^2 = 0.008$). Thus, the DA-SR approach did not benefit the students' English speaking skills in comparison with the CF-SR approach.

Table 3. The analysis of the ANCOVA on the two groups' English speaking skills

Group	*n*	Mean	*SD*	Adjusted mean	Std. error	*F*	η^2
DA-SR group	30	57.50	17.57	54.35	2.22	0.43	0.008
CF-SR group	26	48.54	20.68	52.18	2.39		

Note. Both tests were scored on a scale of 0-80.

5.2. English speaking learning anxiety

As regards the students' English speaking learning anxiety, the paired *t* tests were computed to respectively explore the impacts that the two learning approaches had on the participants. In this study, the English speaking learning anxiety included three dimensions, namely "English class performance anxiety," "lack of self-confidence in English," and "fear of ambiguity in English." As depicted in Table 4, significant differences were verified with $t = 4.98$ ($p < .001$) for the "English class performance anxiety" dimension, and $t = 5.49$ ($p < .001$) for the "fear of ambiguity" dimension, and $t = 3.72$ ($p < .01$) for the total factors in English speaking learning anxiety. This implies that the DA-SR approach can effectively reduce students' perceptions of English speaking learning anxiety, especially the "English class performance anxiety," and the "fear of ambiguity" dimensions. On the other hand, as displayed in Table 5, significant differences were found with $t = 2.49$ ($p < .05$) for the "fear of ambiguity" dimension, and $t = 2.12$ ($p < .05$) for total factors in English speaking learning anxiety, indicating that the CF-SR approach can significantly decrease students' English speaking learning anxiety, especially the "fear of ambiguity" dimension.

Moreover, the one-way ANCOVA was applied to investigate the effects of the different learning approaches on the post-questionnaire ratings of the three dimensions, and the individual pre-questionnaire ratings were utilized as the covariates. To determine whether the adoption of ANCOVA was acceptable, the homogeneity test was executed first. The homogeneity of the regression slopes was confirmed with $F= 0.13$ ($p > .05$) for the "English class performance anxiety" dimension, $F = 2.12$ ($p > .05$) for the "lack of self-confidence" dimension, and $F = 0.27$ ($p > .05$) for the "fear of ambiguity" dimension.

Following that, the ANCOVA was conducted and the results are displayed in Table 6. A significant difference was found in the post-questionnaire ratings of the students' English class performance anxiety ($F = 4.08$, $p < .05$, $\eta^2 = 0.071$), whereas no significant difference was displayed in those of their perceptions of lack of self-confidence ($F = 0.83$, $p > .05$, $\eta^2 = 0.015$), or in their perceptions of fear of ambiguity in English ($F = 2.50$, $p > .05$, $\eta^2 = 0.045$) via precluding the interference from the pre-questionnaire ratings. Furthermore, the adjusted means of the post-questionnaire ratings of the students' English class performance anxiety were 2.07 (Std. error = 0.10) for the DA-SR group, and 2.37 (Std. error = 0.11) for the CF-SR group, describing that the DA-SR approach can significantly reduce students' English class performance anxiety in comparison with the CF-SR approach. According to Cohen's (1988) assertion, the effect size for the different learning approaches was medium ($\eta^2 > 0.059$) for students' English class performance anxiety.

Table 4. The paired *t*-test result of the experimental group's English speaking learning anxiety

Factor	Variable and source	*n*	Mean	*SD*	*t*
English class performance anxiety	Pre-survey	30	2.77	1.02	4.98^{***}
	Post-survey	30	2.15	0.84	
Lack of self-confidence	Pre-survey	30	2.87	1.17	-0.13
	Post-survey	30	2.89	1.24	
Fear of ambiguity	Pre-survey	30	2.67	1.17	5.49^{***}
	Post-survey	30	2.04	0.95	
Total factors	Pre-survey	30	2.77	1.07	3.72^{**}
	Post-survey	30	2.36	0.94	

Note. $^{**}p < .01$; $^{***}p < .001$.

Table 5. The paired *t*-test result of the control group's English speaking learning anxiety

Factor	Variable and source	*n*	Mean	*SD*	*t*
English class performance anxiety	pre-survey	26	2.50	0.83	1.90
	post-survey	26	2.28	0.77	
Lack of self-confidence	pre-survey	26	3.12	1.30	1.29
	post-survey	26	2.78	1.20	
Fear of ambiguity	pre-survey	26	2.72	0.98	2.49^{*}
	post-survey	26	2.34	1.04	
Total factors	pre-survey	26	2.77	0.94	2.12^{*}
	post-survey	26	2.47	0.87	

Note. $^{*}p < .05$.

Table 6. The ANCOVA analysis of the two groups' English speaking learning anxiety

Variable and source	Group	*n*	Mean	*SD*	Adjusted mean	Std. error	*F*	η^2
English class performance anxiety	DA-SR group	30	2.15	0.84	2.07	0.10	4.08^{*}	0.071
	CF-SR group	26	2.28	0.77	2.37	0.11		
Lack of self-confidence	DA-SR group	30	2.89	1.24	2.95	0.19	0.83	0.015
	CF-SR group	26	2.78	1.20	2.71	0.20		
Fear of ambiguity	DA-SR group	30	2.04	0.95	2.05	0.12	2.50	0.045
	CF-SR group	26	2.34	1.04	2.33	0.13		

Note. $^{*}p < .05$.

5.3. Cognitive load

In this study, the cognitive load survey comprised two dimensions, namely "mental effort" and "mental load." The independent *t* tests were utilized to investigate the effects of the different learning approaches on students' intrinsic and extraneous cognitive load.

As regards the mental load dimension (as presented in Table 7), no significant difference was found in the two groups' questionnaire ratings with $t = -1.99$ ($p > .05$), describing that there is no significantly different effect of the two approaches on students' intrinsic cognitive load. On the other hand, in terms of the mental effect dimension, a significant difference was verified between the two groups' mental effort, with $t = -2.17$ ($p < .05$). The means were respectively 2.02 ($SD = 0.92$) and 2.60 ($SD = 1.08$) for the DA-SR group, and for the CF-SR group, revealing that the students who learned with the DA-SR approach were conscious of lower extraneous cognitive load than the ones who learned with the CF-SR approach. Furthermore, all the means of the two groups' questionnaire ratings were below average (Mean = 3), suggesting that all participants perceived low cognitive load during the different learning activities.

Table 7. The *t*-test result of the two groups' cognitive load levels

Variable and source	Group	*n*	Mean	*SD*	*t*
Mental load	DA-SR group	30	1.80	0.78	-1.99
	CF-SR group	26	2.32	1.14	
Mental effort	DA-SR group	30	2.02	0.92	-2.17^{*}
	CF-SR group	26	2.60	1.08	

Note. $^{*}p < .05$.

6. Discussion and conclusions

In this study, a dynamic assessment-based speech recognition approach was implemented to enhance students' English speaking learning. A learning activity was conducted in an elementary school English course. The experimental results revealed that both the experimental group (DA-SR) and the control group (CF-SR) effectively improved their English speaking skills, and decreased their perceptions of English speaking learning anxiety. Moreover, the DA-SR approach successfully reduced the students' English class performance anxiety and extraneous cognitive load in comparison with the CF-SR approach.

Speaking anxiety is regarded as a crucial factor that could affect students' speaking competence (Chen & Hwang, 2020), yet it is argued that dealing with CF could be stressful, resulting in great learning anxiety (Bodnar et al., 2017). Both the groups learning with the two different speech recognition systems significantly reduced their perceptions of English speaking learning anxiety. Such a finding corresponds to what has been depicted by Rassaei (2019), who emphasized the significance of integrating proper CF into technology-enhanced speaking instruction for EFL learning. This also confirms what has been portrayed by several reports, namely that a speech recognition approach, if properly designed, is capable of reducing learners' English speaking anxiety (de Vries et al., 2015; Wang & Young, 2014).

With respect to the three dimensions of students' English speaking learning anxiety, it is evidenced that both speech recognition approaches effectively lower students' English speaking learning anxiety for the "fear of ambiguity" dimension. This result is similar to the view asserted by Li (2010) and Li and Vuono (2019), who stated that using CF can reply to the appropriateness of a learner's production or comprehension. This could be the reason why the two speech recognition approaches with different types of CF are of great benefit in terms of reducing students' fear of ambiguity in English speaking. Moreover, only the DA-SR approach significantly decreased the students' perceptions of English class performance anxiety. It is asserted that DA is capable of adapting to individual learning preferences with fine-tuned interventions (Bakhoda & Shabani, 2019). This could explain the effect that the students who learned with the DA-SR approach revealed a lower degree of English class performance anxiety than those who learned with the CF-SR approach.

By way of illustration, Penning de Vries et al. (2020) asserted the significance of taking into consideration working memory capacity when designing CF in the English speaking learning activity. Both groups, which adopted the two speech recognition systems with the different types of CF, perceived low cognitive load. This could be a good illustration for combining CF with a speech recognition system in the English speaking learning activity. Moreover, the DA-SR approach significantly reduced the students' extraneous cognitive load in comparison with the CF-SR approach. DA and the dialogic mediation can stimulate students' reflection, and accordingly less feedback and explanations are demanded (Andujar, 2020). This could explain why the students who learned with the DA-SR approach were conscious of lower extraneous cognitive load. It is also described that excess information processing during the learning process may lead to more extraneous cognitive load (Leahy, & Sweller, 2016; Wu et al., 2018), indicating the notable value of a well-designed DA-SR approach in English speaking learning activities.

All participants in this experiment significantly enhanced their English speaking skills, implying the importance of adequately integrating CF and a speech recognition system. This accords with the view of Couper (2019) and Rassaei (2019), who described the impacts of oral CF on students' speaking abilities. As mentioned above, the CF-SR approach successfully promoted students' English speaking skills, and reduced their learning anxiety by means of properly integrating CF into a speech recognition system. This study further demonstrated that the DA-SR approach can lower students' English class performance anxiety, and extraneous cognitive load. This also supports the notion revealed by several studies (e.g., de Vries et al., 2015; Young & Wang, 2014), which emphasized that ASR-based learning systems should provide CF to promote learners' speaking skill, and to reduce their cognitive load.

This study designed the DA-SR approach for English speaking, and effectively promoted students' speaking learning effectiveness. It could be a valuable reference for designing English speaking activities in EFL learning environments. Nevertheless, neither group reduced their learning anxiety related to their self-confidence in English speaking. It is suggested that more different types of CF could be adopted in further studies regarding ASR-based learning systems. It is also worth investigating the effects of using an ASR technology in different learning contexts on students' English speaking, such as game-based learning.

Acknowledgement

This study is supported in part by the Ministry of Science and Technology of the Republic of China under Contract Numbers 109-2511-H-142-011.

References

Abdullah, M. Y., Hussin, S., & Ismail, K. (2019). Investigating the effects of the flipped classroom model on Omani EFL learners' motivation level in English speaking performance. *Education and Information Technologies, 24*(5), 2975-2995.

Ahn, T. Y., & Lee, S. M. (2016). User experience of a mobile speaking application with automatic speech recognition for EFL learning. *British Journal of Educational Technology, 47*(4), 778-786.

Allal, L., & Ducrey, G. P. (2000). Assessment of-or in-the zone of proximal development. *Learning and Instruction, 10*(2), 137-152.

Andujar, A. (2020). Mobile-mediated dynamic assessment: A New perspective for second language development. *ReCALL, 32*(2), 178-194.

Antón, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals, 42*(3), 576-598.

Bakhoda, I., & Shabani, K. (2019). Bringing L2 learners' learning preferences in the mediating process through computerized dynamic assessment. *Computer Assisted Language Learning, 32*(3), 210-236.

Bodnar, S., Cucchiarini, C., Penning de Vries, B., Strik, H., & van Hout, R. (2017). Learner affect in computerised L2 oral grammar practice with corrective feedback. *Computer Assisted Language Learning, 30*(3-4), 223-246.

Cavus, N., & Ibrahim, D. (2017). Learning English using children's stories in mobile devices. *British Journal of Educational Technology, 48*(2), 625-641.

Chen, C. H. (2020). AR videos as scaffolding to foster students' learning achievements and motivation in EFL learning. *British Journal of Educational Technology, 51*(3), 657-672.

Chen, H. H. J. (2011). Developing and evaluating an oral skills training website supported by automatic speech recognition technology. *ReCALL, 23*(1), 59-78.

Chen, M. R. A., & Hwang, G. J. (2020). Effects of a concept mapping-based flipped learning approach on EFL students' English speaking performance, critical thinking awareness and speaking anxiety. *British Journal of Educational Technology, 51*(3), 817-834.

Chen, X., Xie, H., & Hwang, G. J. (2020a). A Multi-perspective study on Artificial Intelligence in Education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence, 1,* 100005. https://doi.org/10.1016/j.caeai.2020.100005

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020b). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education, 146,* 103751. https://doi.org/10.1016/j.compedu.2019.103751

Cho, E., Compton, D. L., & Josol, C. K. (2020). Dynamic assessment as a screening tool for early identification of reading disabilities: A Latent change score approach. *Reading and Writing, 33*(3), 719-739.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Couper, G. (2019). Teachers' cognitions of corrective feedback on pronunciation: Their beliefs, perceptions and practices. *System, 84*, 41-52.

Crescenzi-Lanna, L. (2020). Multimodal learning analytics research with young children: A Systematic review. *British Journal of Educational Technology, 51*(5), 1485-1504.

de Vries, B. P., Cucchiarini, C., Bodnar, S., Strik, H., & van Hout, R. (2015). Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning, 28*(6), 550-576.

Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning, 32*(5-6), 527-555.

Fathi, J., & Ebadi, S. (2020). Exploring EFL pre-service teachers' adoption of technology in a CALL program: Obstacles, motivators, and maintenance. *Education and Information Technologies, 25*, 3897-3917.

Fu, Q. K., Lin, C. J., Hwang, G. J., & Zhang, L. (2019). Impacts of a mind mapping-based contextual gaming approach on EFL students' writing performance, learning perceptions and generative uses in an English course. *Computers & Education, 137*, 59-77.

Gan, Z. (2013). Understanding English speaking difficulties: An Investigation of two Chinese populations. *Journal of Multilingual and Multicultural Development, 34*(3), 231-248.

Herazo, J. D., Davin, K. J., & Sagre, A. (2019). L2 dynamic assessment: An Activity theory perspective. *The Modern Language Journal, 103*(2), 443-458.

Hsu, L. (2016). An Empirical examination of EFL learners' perceptual learning styles and acceptance of ASR-based computer-assisted pronunciation training. *Computer Assisted Language Learning, 29*(5), 881-900.

Huang, Y. M., Shadiev, R., & Hwang, W. Y. (2016). Investigating the effectiveness of speech-to-text recognition applications on learning performance and cognitive load. *Computers & Education, 101*, 15-28.

Hwang, G. J., & Fu, Q. K. (2019). Trends in the research design and application of mobile language learning: A Review of 2007–2016 publications in selected SSCI journals. *Interactive Learning Environments, 27*(4), 567-581.

Hwang, G. J., Sung, H. Y., Chang, S. C., & Huang, X. C. (2020a). A Fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Computers and Education: Artificial Intelligence, 1*, 100003. https://doi.org/10.1016/j.caeai.2020.100003

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020b). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A Concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education, 69*, 121-130.

Kuru Gönen, S. İ. (2019). A Qualitative study on a situated experience of technology integration: Reflections from pre-service teachers and students. *Computer Assisted Language Learning, 32*(3), 163-189.

Lai, A. F., Chen, C. H., & Lee, G. Y. (2019). An Augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology, 50*(1), 232-247.

Leahy, W., & Sweller, J. (2016). Cognitive load theory and the effects of transient information on the modality effect. *Instructional Science, 44*(1), 107-123.

Li, S. (2010). The Effectiveness of corrective feedback in SLA: A Meta-analysis. *Language Learning, 60*(2), 309-365.

Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in System. *System, 84*, 93-109.

Liu, M., & Jackson, J. (2008). An Exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal, 92*(1), 71-86.

Luo, B. (2016). Evaluating a computer-assisted pronunciation training (CAPT) technique for efficient classroom instruction. *Computer Assisted Language Learning, 29*(3), 451-476.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 20*, 37-66.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching, 46*(1), 1-40.

Nguyen, T. H., Hwang, W. Y., Pham, X. L., & Ma, Z. H. (2018). User-oriented EFL speaking through application and exercise: Instant speech translation and shadowing in authentic context. *Educational Technology & Society, 21*(4), 129-142.

Paas, F. G. W. C, Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71.

Paas, F., & van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*(4), 351-371.

Penning de Vries, B. W., Cucchiarini, C., Strik, H., & Van Hout, R. (2020). Spoken grammar practice in CALL: The Effect of corrective feedback and education level in adult L2 learning. *Language Teaching Research, 24*(5), 714-735.

Rassaei, E. (2019). Computer-mediated text-based and audio-based corrective feedback, perceptual style and L2 development. *System, 82*, 97-110.

Rezaee, A. A., Alavi, S. M., & Razzaghifard, P. (2019). The Impact of mobile-based dynamic assessment on improving EFL oral accuracy. *Education and Information Technologies, 24*(5), 3091-3105.

Rezaee, A. A., Alavi, S. M., & Razzaghifard, P. (2020). Mobile-based dynamic assessment and the development of EFL students' oral fluency. *International Journal of Mobile Learning and Organisation, 14*(4), 511-532.

Shadiev, R., Hwang, W. Y., Chen, N. S., & Huang, Y. M. (2014). Review of speech-to-text recognition technology for enhancing learning. *Educational Technology & Society, 17*(4), 65-84.

Shadiev, R., Wu, T. T., Sun, A., & Huang, Y. M. (2018). Applications of speech-to-text recognition and computer-aided translation for facilitating cross-cultural learning through a learning activity: Issues and their solutions. *Educational Technology Research and Development, 66*(1), 191-214.

Swanson, H. L., & Lussier, C. M. (2001). A Selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research, 71*(2), 321-363.

Thompson, A. S., & Lee, J. (2013). Anxiety and EFL: Does multilingualism matter? *International Journal of Bilingual Education and Bilingualism, 16*(6), 730-749.

Wang, J. R., & Chen, S. F. (2016). Development and validation of an online dynamic assessment for raising students' comprehension of science text. *International Journal of Science and Mathematics Education, 14*(3), 373-389.

Wang, Y. H., & Young, S. S. C. (2014). A Study of the design and implementation of the ASR-based iCASL system with corrective feedback to facilitate English learning. *Educational Technology & Society, 17*(2), 219-233.

Wu, H. M., Kuo, B. C., & Wang, S. C. (2017). Computerized dynamic adaptive tests with immediately individualized feedback for primary school mathematics learning. *Educational Technology & Society, 20*(1), 61-72.

Wu, P. H., Hwang, G. J., Yang, M. L., & Chen, C. H. (2018). Impacts of integrating the repertory grid into an augmented reality-based learning design on students' learning achievements, cognitive load and degree of satisfaction. *Interactive Learning Environments, 26*(2), 221-234.

Yang, H., & Meinel, C. (2014). Content based lecture video retrieval using speech and video text information. *IEEE Transactions on Learning Technologies, 7*(2), 142-154.

Yang, Y. F., & Kuo, N. C. (2020). New teaching strategies from student teachers' pedagogical conceptual change in CALL. *System, 90*, 102218. https://doi.org/10.1016/j.system.2020.102218

Young, S. S. C., & Wang, Y. H. (2014). The Game embedded CALL system to facilitate English vocabulary acquisition and pronunciation. *Educational Technology & Society, 17*(3), 239-251.

Zhang, R., & Zou, D. (2020). Influential factors of working adults' perceptions of mobile-assisted vocabulary learning with multimedia annotations. *International Journal of Mobile Learning and Organisation, 14*(4), 533-548.

Zhang, Y., & Liu, L. (2018). Using computer speech recognition technology to evaluate spoken English. *Educational Sciences: Theory & Practice, 18*(5), 341-1350.