Computers & Education 101 (2016) 15–28

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Investigating the effectiveness of speech-to-text recognition applications on learning performance and cognitive load

Yueh-Min Huang [a], Rustam Shadiev [a, *], Wu-Yuin Hwang [b]

[a] Department of Engineering Science, National Cheng Kung University, No.1, University Road, Tainan 70101, Taiwan
[b] Graduate Institute of Network Learning Technology, National Central University, No.300, Jhongda Rd., Jhongli 32001, Taiwan

A R T I C L E I N F O

Article history:
Received 23 October 2015
Received in revised form 17 May 2016
Accepted 25 May 2016
Available online 27 May 2016

Keywords:
Improving classroom teaching
Pedagogical issues
Post-secondary education

A B S T R A C T

This study explores the effectiveness of applying speech-to-text recognition (STR) technology during lectures in English on learning performance and the cognitive load of non-native English speaking students. Furthermore, the study also explores the usefulness of texts generated using STR for students with different levels of English as foreign language (EFL) ability during lectures of varying difficulty levels. Two lectures, one with intermediate difficulty level content and the other advanced, were administered, and STR was adopted to aid student learning. The results of this study show that the students who used STR-generated texts outperformed the students who did not. Furthermore, lectures in English caused less cognitive load for low ability EFL students when they used STR-texts. According to the students, the STR-texts were useful for following the instructor, confirming content, clarifying vocabulary, and making up missed information. It was found that STR-texts were used by low EFL ability students during both lectures whereas high EFL ability students used STR-texts during the lecture at the advanced level and only some of high EFL ability students used them during the intermediate lecture. Based on these results, several suggestions and implications for teaching and research community are proposed.

1. Introduction

1.1. Research gaps

English is presently the most widely used language in the academic world (Shadiev, Hwang, Huang, & Liu, 2016) and is used as the main communication language in most academic events around the world. However, recent evidence suggests that some participating non-native speakers at academic events are unable to fully comprehend communication content (Camiciottoli, 2005; Shadiev, Hwang, & Huang, 2013). Miller (2007) argued that many non-native English speaking participants flounder during speeches delivered in English, and they have to exert extra effort to comprehend what they are hearing. Other studies have demonstrated that non-native English speaking participants are usually silent and seldom interact during academic events carried out in English (Bain, Basson, Faisman, & Kanevsky, 2005; Barnes & Lock, 2010; Camiciottoli, 2005; Kramer, Walker, & Brill, 2007; Miller, 2007; Poulsen, Hastings, & Allbritton, 2007). To address this issue, application of

* Corresponding author.
E-mail address: rustamsh@gmail.com (R. Shadiev).

http://dx.doi.org/10.1016/j.compedu.2016.05.011

speech-to-text recognition (STR) technology is suggested as a potentially reliable and valuable tool for non-native English speaking students to help them better comprehend English speeches (Ryba, McIvor, Shakir, & Paez, 2006). STR synchronously transcribes text streams from a lecturer's speech input onto a whiteboard or computer screens. Thus, STR-generated texts can enhance the comprehension of non-native English speaking students during academic events held in English.

The pedagogical usefulness of STR-texts to enhance learning has been emphasized in several studies. According to Shadiev, Hwang, Chen, and Huang (2014), STR can be applied to enhance the learning of students with cognitive or physical disabilities, online students, and students learning in physical classrooms. Table 1 presents STR-related studies.

Firstly, previous studies have provided evidence on the effectiveness of STR-text on learning based on statistical analysis. However, the number of subjects involved in these studies has been consistently insufficient (Table 1), for example, $n = 21$ in Huang, Liu, Shadiev, Shen, and Hwang (2015), $n = 44$ in Hwang, Shadiev, Kuo, and Chen (2012), $n = 37$ in Kuo, Shadiev, Hwang, and Chen (2012) and $n = 9$ in Ranchal et al. (2013). Creswell (2014) and Fraenkel, Wallen, and Hyun (2014) warned that an experimental study needs sufficient sample size for data collection to obtain adequate power and to meet the assumptions required of specific statistical tests used in analysis. They argued that the minimum acceptable sample size must be at least 30 in each group for experimental and causal comparative studies since data obtained from a smaller number of subjects may give inaccurate estimates and provide meaningless results.

Secondly, Huang et al. (2015) assigned students who used STR-texts during lectures into low ability ($n = 7$) and high ability ($n = 7$) groups and then compared their pre-test and post-test scores. In the study of Ranchal et al. (2013), nine students attended lectures with STR-texts for the first six weeks and without STR-texts for the next six weeks. The scores of the nine students for each six-week period were then compared. That is, in Huang et al. (2015), the effectiveness of the STR-texts on learning was not tested by comparing the scores of students who used STR-texts (a treatment group) with students who did not use STR-texts (a control group) but by comparing the scores of students with different learning abilities. Furthermore, in Ranchal et al. (2013), the scores of the same group of students were compared after attending lectures with STR-texts and after attending lectures without STR-texts. Such methodologies conflict with experimental design principles and may misrepresent the effects of an intervention (Creswell, 2014; Fraenkel et al., 2014).

Thirdly, a pretest-and-posttest design used to compare prior knowledge and learning achievement of students who learned with STR (an experimental group) with those without it (a control group) in a synchronous online learning environment was adopted in Hwang et al. (2012) and Kuo et al. (2012). One may argue that the student-lecture interactions are

Table 1
STR-related studies.

Reference	Participants number	Native language	Language of instruction	Design	Research focus
Huang et al. (2015)	21	Chinese	English	1. Eye-tracking technique 2. Questionnaire 3. Compare pre-test and post-test scores of low ability ($n = 7$) and high ability ($n = 7$) students	1. Visual attention and learning behavior to use STR-texts 2. Students perceptions 3. The effectiveness of STR-texts on learning
Hwang et al. (2012)	Control group: 19 Experimental group: 25	Chinese	Chinese	1. Compare pre-test and post-test results of control and experimental students	1. The effectiveness of using/not using STR-texts on learning
Kuo et al. (2012)	Control group: 21 Experimental group: 16	Chinese	Chinese	1. Compare pre-test and post-test results of control and experimental students	1. The effectiveness of using/not using STR-texts on learning
Leitch (2008)	44	English	English	1. Questionnaire survey	1.Teachers and students experience with STR 2. Teachers and students perceptions toward STR
Ranchal et al. (2013)	9	English	English	1. Compare quiz scores of students when they learned with notes and without notes	1.The effectiveness of learning with/without notes on learning performance
Ryba et al. (2006)	160	English: 79 Non-English: 81	English	1. Email responses 2. WebCT discussion forum 3. Questionnaire	1. Application of STR to assist native and non-native English speaking students
Shadiev et al. (2016)	9	Chinese	English	1. Interview surveys 2. Content analysis of student essays	1. Student perceptions toward usefulness of STR 2. Learning strategies and behavior to use STR-texts
Wald and Bain (2008)	NA	English	English	1. Questionnaire survey	1. Teachers and students experience with STR 2. Teachers and students perceptions toward STR

basically the same (lectures including PowerPoint slides and video with/without STR-texts are viewed on computer screens) when lectures are delivered through computer screens in an online or physical classroom. However, they are different since online students use STR-texts for reasons that are not applicable to students in physical classrooms (e.g. online students cannot hear their instructor due to low internet bandwidth and are more vulnerable to distractions than students in physical classrooms). Furthermore, students in Hwang et al. (2012) and Kuo et al. (2012) were native speakers of Chinese listening to Chinese lectures. On the other hand, studies of STR with students for whom lectures were carried out in a foreign language used qualitative data such as a questionnaire survey (Leitch, 2008; Ryba et al., 2006; Wald & Bain, 2008), email responses, WebCT discussion forums (Ryba et al., 2006), and interviews (Shadiev et al., 2016) to determine the effectiveness of STR. The findings of these studies are based on subjective rather than empirical evidence obtained from statistical analysis and are therefore liable to the failings of this kind of research.

Fourthly, most studies on STR have focused on the effectiveness of STR-texts on learning and students' experiences with STR (Huang et al., 2015; Hwang et al., 2012; Kuo et al., 2012; Leitch, 2008; Ranchal et al., 2013; Ryba et al., 2006; Shadiev et al., 2016; Wald & Bain, 2008). In spite of the advantages of STR-texts, it is still unclear whether using them in lectures in a foreign language impose more or less cognitive load compared to lectures without them or whether this use affects students with heterogeneous language abilities during lectures when the learning content is on different levels of difficulty (Huang et al., 2015).

1.2. Related theories

Two cognitive theories, *cognitive theory of multimedia learning* and *cognitive load theory*, may help explain how students' learning may differ when learning with or without STR and how students with heterogeneous language abilities may experience different cognitive load during lectures with the learning content is offered at different levels of difficulty, in turn, resulting in different learning outcomes.

The cognitive theory of multimedia learning was proposed by Mayer (2009) to explain how a learner learns using content presented in different forms (i.e. multimedia). According to this theory, visual and verbal materials are processed in different parts of the brain. The visual channel takes input initially from the eyes and ultimately produces pictorial representations; the verbal channel takes input initially from the ears and ultimately produces verbal representations. During speech, a learner pays attention to an auditory message, parses speech, segments it into words that are kept in verbal working memory, and then transforms the words into verbal mental representations. After this process, the connections are mentally built to organize words into cause-and-effect chains. Similarly, a learner pays attention to visual messages, selects images and holds them in visual working memory. A learner also mentally builds connections that organize images into cause-and-effect chains. Finally, the visual model, verbal mental model, and prior knowledge are merged through constructing referential connections among them.

Cognitive load theory is a highly effective guide for the design of multimedia learning (Mayer & Moreno, 2003; Paas, Renkl, & Sweller, 2003). Three types of cognitive load are distinguished in the literature: intrinsic, extraneous, and germane (Brunken, Plass, & Leutner, 2003; Sweller, Van Merrienboer, & Paas, 1998). Intrinsic load is determined by the inherent nature of the learning material, learners' expertise and the interaction between them—that is, the amount of information units that a learner needs to hold in working memory to comprehend the information being acquired. It has been argued that intrinsic load is not affected by the instructional design but only by the learning material. Therefore, to avoid intrinsic load overloading, instructors need to adjust learning materials to meet the learners' levels of expertise. The term "extraneous load" refers to the cognitive load caused by the format and the manner in which information is presented as well as by the working memory requirements required to do the instructional activities. An extraneous load can be imposed by improper instructional design. Thus, to keep the extraneous load from becoming excessive, instructors need to organize and present learning information and carry out learning activities appropriately. A germane load is determined by learners' efforts to process and comprehend learning material. This load is also associated with motivation and interest. A germane load is induced by appropriate instructional design and can enhance learning.

Cognitive load theory suggests that working memory limits learners' cognitive capacity to accommodate demands imposed by learning tasks (Paas et al., 2003). This is especially important when the lecture is delivered in a language that is not native to learners. Keysar, Hayakawa, and An (2012) claimed that some information processed during learning in a foreign language is complex and imposes a heavy cognitive load on working memory. For example, attending a lecture in a foreign language involves high elements interactivity (Plass, Chun, Mayer, & Leutner, 2003). To comprehend the content, learners must receive and retain lecture information in working memory and integrate it with what follows, all the while continually adjusting their understanding with prior knowledge (Chen & Chang, 2009). Mayer and Moreno (2003) and Paas et al. (2003) warned that learning performance can be negatively affected when cognitive load exceeds the limit of cognitive capacity. Thus, the issue of how to reduce cognitive capacity overload while encouraging knowledge construction should be considered when attempting to improve instructional designs. Therefore, several approaches, including application of STR technology, have been proposed in the literature to help students during lectures in a foreign language (Ryba et al., 2006).

Mayer and Moreno (1998) claimed that the same information presented in auditory and written forms makes it redundant and gives rise to a split-attention effect and increased cognitive load. For example, the redundancy effect takes place in lectures with STR support since the same information in verbal (i.e. a lecturer's speech) and visual (STR-texts) forms is presented to learners simultaneously. However, some exceptions have been suggested regarding the redundancy principle.

For example, Clark and Mayer (2011) argued that there are special situations in which redundant visual text is acceptable and even necessary, especially when the verbal information is difficult to understand, such as when it is presented in a foreign language. In these situations, visual text is helpful. On top of this, the expertise reversal principle states that instructional techniques that are highly effective with novice learners may not be effective and may even have negative consequences, such as increased cognitive load, when used with more knowledgeable learners (Kalyuga, 2014). For example, low language ability learners need learning information presented both in visual and verbal forms to understand it better because they have low language ability and lack prior knowledge. On the contrary, experts already have some prior knowledge, and information in two modalities could become not only redundant but even counterproductive since processing it may require additional cognitive resources.

In our study, non-native English speaking students attended lectures in English. We also displayed texts generated by STR from the speech input for students during the lectures. This approach allowed students to receive information in two forms: verbal (i.e. speech of the instructor) and visual (i.e. STR-texts). According to the cognitive theory of multimedia learning given above, presenting verbal and visual content simultaneously contributes to better processing of information. That is, the verbal and visual processing systems are coordinated to help process and integrate learning content presented in two forms. When a student builds referential connections between created models and prior knowledge, both visual and verbal mental models can be accessed. We assume that presenting STR-texts during lectures simultaneously can be very useful so that students may learn content better and as a result, may outperform those who study without STR support.

The following studies tested our assumption (see also Table 1); however, there are some research gaps still existing that need attention (please refer to our discussion in earlier sections related to the validity of empirical results regarding the effectiveness of STR-texts on learning). Leitch (2008) and Wald and Bain (2008) applied IBM ViaVoice STR to assist students with hearing impairments to listen to lectures. STR-texts were displayed to students and could be used to take notes and verify and clarify speech content. Leitch (2008) and Wald and Bain (2008) conducted surveys, and their results indicated that STR-texts help students to understand lectures better. Hwang et al. (2012) and Kuo et al. (2012) adopted the STR provided by the Microsoft Operating System to support online teaching and learning activities in a synchronous learning environment. Online students in the studies of Hwang et al. (2012) and Kuo et al. (2012) could simultaneously listen to a speaker and read the transcripts of the lecture. Furthermore, STR-texts could be retrieved for post-lecture study. Hwang et al. (2012) and Kuo et al. (2012) carried out experiments, and their results demonstrated that online students who used transcripts performed better than online students who did not use transcripts on assignments and post-tests. Furthermore, online students positively perceived STR-texts as useful for learning and stated their willingness to use STR-texts for learning in the future (Hwang et al., 2012; Kuo et al., 2012). Students assert that STR-texts are useful for understanding the speaker, catching up on missed information, taking notes, and completing homework. IBM ViaScribe STR was employed to assist university students with life and social science course lectures (Ranchal et al., 2013) and information systems (Ryba et al., 2006). Ranchal et al. (2013) and Ryba et al. (2006) provided students with two distinct methods for acquiring the STR-mediated lectures. One is real-time captioning, in which verbal information is processed into textual captions with the help of the STR and streamed directly onto students' computers screens. The other is post-lecture transcription, i.e. lecture transcripts in which STR-recognition errors had already been corrected and made available to students after class. Ryba et al. (2006) explored both students' perceptions of using STR-texts and the extent that they actually make use of them. The results showed that more than 30% of the students used STR-texts for learning while more than 40% of students tended to use STR-texts (Ryba et al., 2006). Students have mentioned that STR-texts help them to understand lectures, confirm what they miss in lectures, and take notes. Ranchal et al. (2013) evaluated the effectiveness of real-time lecture transcriptions and post-lecture transcriptions. Their results showed that students benefit from both methods: real-time lecture transcriptions allow students to pay attention to the instructor instead of taking class notes; post lecture transcriptions enable students to study lecture content after class and to perform better on quizzes.

A significant body of literature has looked at how STR application supports the learning of non-native English speaking students during lectures in English (Table 1). Ryba et al. (2006) applied IBM ViaScribe STR for this purpose during three 2-h lectures delivered in English by a non-native speaker. Spoken lectures were transcribed into text using STR and displayed on a large screen in front of the lecture hall. Ryba et al. (2006) surveyed students afterwards and found that STR was beneficial to some. Non-native speaker students found the STR-texts useful for following the instructor, clarifying lecture content, and enhancing their understanding. Shadiev et al. (2016) adopted an STR system provided by the Microsoft Operating System in a graduate seminar program on advanced learning technologies. The seminar was conducted in English for nine non-native speaking participants. Their results showed that most participants had used more than half of the STR-texts to complete summary writing tasks. A participant survey showed that they made good use of nineteen learning strategies related to the use of STR-texts in order to understand presented topics, answer questions, and complete assignments. More importantly, the survey results showed that STR-texts facilitated and enhanced learning performance. It was also found that most participants perceived STR-texts to be useful for learning both during and after the seminar. Huang et al. (2015) used Window's STR to transcribe two English lectures to assist non-native English speaking participants with comprehending lectures and investigated the attention participants paid to the STR-texts and how STR-texts influence learning achievement. Their results showed STR-texts to be useful during the lectures as learning aids and that participants relied on STR-texts more than on the video of the instructor or the slides.

Regarding cognitive load, based on related studies (Mayer & Moreno, 2003; Paas et al., 2003), we assume here that lectures in foreign language will result in a heavy cognitive load on the working memory of students due to the complex information processes involved (Chen & Chang, 2009; Keysar et al., 2012; Plass et al., 2003). However, if students receive learning information in both verbal and visual forms, STR-texts will be helpful with regard to diminishing information processing demands. That is, both verbal and visual working memories will be employed during information processing instead of verbal working memory only. In this case, neither type of working memory will be overloaded, and learners' cognitive capacity to accommodate demands imposed by learning tasks will be increased (Clark & Mayer, 2011).

We also assume that the redundancy and expertise reversal effects will take place during lectures with STR support when lecture content difficulty level matches the language ability of the students (Clark & Mayer, 2011; Kalyuga, 2014). That is, when the same information is presented to students during a lecture both verbally (i.e. a lecturer's speech) and visually (STR-texts) the students will be cognitively overloaded. When students are processing complex information during a lecture (Keysar et al., 2012; Plass et al., 2003), which challenges their language abilities, we assume that STR-texts will be beneficial for learning. That is, no redundancy and expertise reversal effects will take place when the same information is presented verbally and visually, but rather each modality will contribute to students' information processing. These two assumptions have not been tested previously. It is therefore still unclear whether the use of STR-texts in lectures in a foreign language imposes more or less cognitive load compared to lectures without STR-texts or whether the use of STR-texts affects students with heterogeneous language ability during lectures when the learning content is at different levels of difficulty. Therefore, these assumptions are tested in our study.

2. Research motivation and questions

In spite of the advantages of STR-texts, the present study has uncovered two potential issues that have not yet been addressed appropriately in the related research. The first issue relates to the validity of empirical results from earlier studies regarding the effectiveness of STR-texts on learning. The second issue relates to cognitive load, which has so far been overlooked in most STR-related studies. In this study, an attempt has been made to remedy these problems. In addition, we aimed to test our three assumptions: (1) presenting STR-texts during lectures can be useful to improve learning performance; (2) lectures in a foreign language result in a heavy cognitive load that can be managed if STR-texts are presented during lectures; (3) students will be cognitively overloaded during lectures in a foreign language with STR-texts if lecture content difficulty level matches their language ability due to the redundancy and expertise reversal effects, but if lecture content is challenging to students' language abilities, they will not be cognitively overloaded. Furthermore, we also aim to investigate students' perceptions of and behavioral intentions toward use of STR-texts for learning. Therefore, this study addresses the following research questions: (1) Do students who use STR-texts have the same learning achievement and cognitive load as those who do not use STR-texts? (2) How does the ability of EFL students affect the benefits of STR-texts on learning achievement and cognitive load? (3) What are perceptions and behavioral intentions of students towards STR-texts?

3. Method

3.1. Participants and experimental procedure

Sixty students from a national university in Taiwan participated in this study. Table 2 shows the students' gender, age, major field of study, and degree distributions.

The experimental procedure for this study is shown in Fig. 1. Before the experiment, fifty students were randomly assigned into either a control or an experimental group. The remaining ten students were assigned into either a control or an experimental group following a matched subjects design in order to match the two groups in terms of gender, age, department, and study for the degree as closely as possible so there would be no significant variance in the variables for the two groups. At the beginning of the experiment, the researchers collected the students' demographic information and

Table 2
Participant demographic profiles.

Demographic characteristics	Control group		Experimental group	
	Frequency	Percentage	Frequency	Percentage
Gender				
Female	16	53.33	16	53.33
Male	14	46.67	14	46.67
Age (years old)				
18-24	30	100	28	93.33
>24	0	0.00	2	6.67
Department				
Applied science	18	60.00	19	63.33
Social Science	12	40.00	11	36.67
Study for the degree				
Undergraduate	24	80.00	25	83.33
Graduate	6	20.00	5	16.67

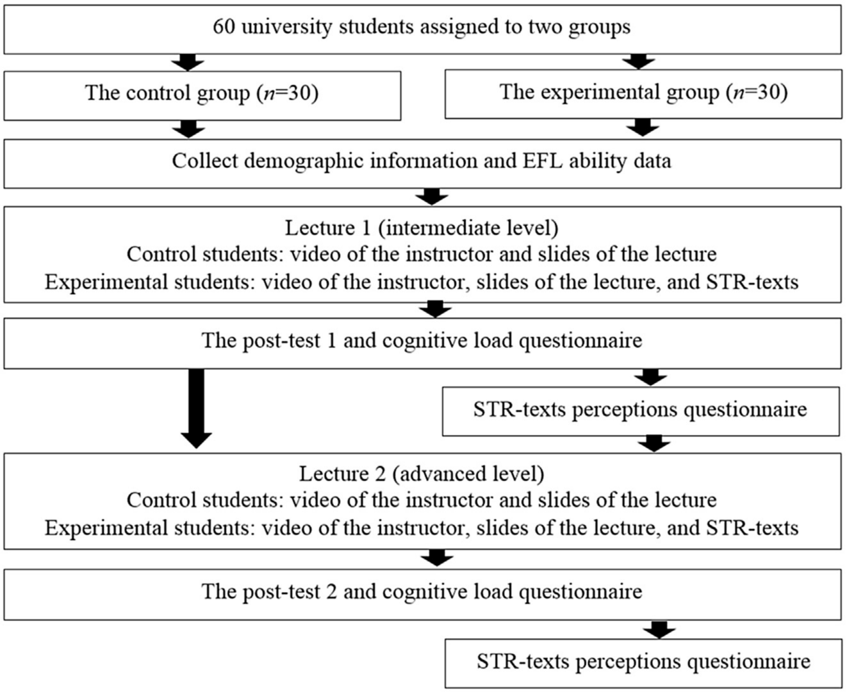

Fig. 1. The experimental procedure.

required them to indicate their English ability by showing their currently-holding certificates of officially-recognized English tests, such as Test of English for International Communication (TOEIC). After that, students attended two lectures given in English by the same instructor. The first lecture "Photography" was at the intermediate level, and the second one "From matchmakers to dating services" was advanced. Both lectures were delivered to students through computer screens. A screenshot of the first video lecture for students in the control and experimental groups is shown as Fig. 2. Students in the control group could see an instructor video and lecture slides, while those in the experimental group could see, in addition to these, STR-generated texts. Of particular interest to this study was providing students with STR-texts to aid their learning of general subjects lectured in a foreign language. That is, English as a foreign language was merely the language of instruction and not the subject the students were learning. The post-test was carried out after each lecture. Finally, the questionnaire survey of students' perceptions and cognitive load was given, and one-on-one semi structured interviews were also conducted.

3.2. Instruments

The students' TOEIC scores were used as the pre-test scores indicating their prior knowledge. Post-test scores were used to assess their learning outcomes. The post-test included five items related to lecture information recognition, two items related

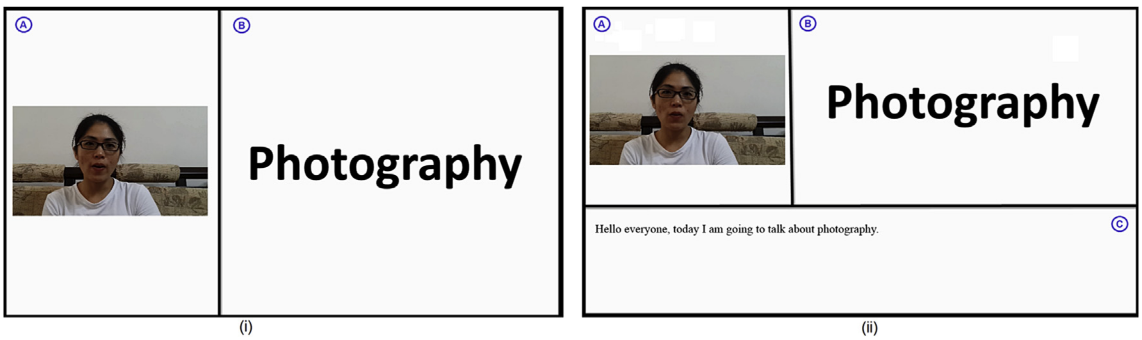

Fig. 2. The first lecture video for the control (i) and experimental (ii) groups: A-video of the instructor, B- the lecture slides, and C- STR-text.

to lecture information recall, and one item measuring their understanding of the lecture. The following are three sample items to measure recognition (1), recall (2), and understanding (3) of learning content after the first lecture "Photography":

1. One rule for taking good pictures is: A) Mistakes can be corrected if you are not afraid. B) Do not make mistakes. C) Making mistakes is normal. D) Don't be afraid to take pictures. Just do your best.
2. What are the other rules?
3. Please summarize the content of the lecture.

Each item related to information recognition was scored as "zero" (incorrect) or "one" (correct). Information recall was tested through open ended questions and summary writing, and students' answers were coded by using a sentence as a coding unit and then scored on a five-point scale (with five as the highest score). The maximum score for the post-test was twenty. Answers to post-test items were scored by three raters, and major differences in assessment were resolved through discussion. The inter-rater reliability of the post-test was evaluated using Cohen's kappa. The analysis results exceeded 0.90, indicating high reliability.

Student cognitive load was measured with a questionnaire designed following general recommendations from previous related studies (Huang, Huang, Liu, & Tsai, 2013; Sweller, Merrienboer, & Paas, 1998). Two items were included in the questionnaire. Item 1, "It was easy to learn the learning material," measured cognitive load, and item 2, "It did not require a lot of mental effort to learn the learning material," measured mental effort. All of the students were asked to respond to the questionnaire, and 60 valid answer sheets were obtained. Responses to the items were scored using a five-point Likert scale, anchored by the end-points "strongly agree" (1) and "strongly disagree" (5). The internal consistency of the survey was tested by employing Cronbach's α; the values exceeded 0.90, demonstrating high item reliability.

For this study, we triangulated different data sources to ground findings related to the effectiveness of STR-texts for learning. That is, the results of the learning performance outcomes and their comparison between the control and experimental groups were supported by the questionnaire survey and interview results.

The perceptions questionnaire survey investigated students' perceptions of the usefulness of STR-texts for learning, as well as their interest in using STR-texts for learning in the future. The design of this questionnaire was informed by previous related studies (Hwang et al., 2012; Kuo et al., 2012) and included the following nine items (items 1–6 related to perceptions and items 7–9 related to behavioral intentions):

1. STR-texts improve my understanding of a lecture.
2. STR-texts increase my productivity during a lecture.
3. STR-texts enhance my learning effectiveness during a lecture.
4. STR-texts improve my learning performance during a lecture.
5. STR-texts help me to accomplish a learning task more quickly.
6. Overall, I found STR-texts to be useful during a lecture.
7. I intend to continue using STR-texts for learning in the future.
8. I plan to use STR-texts for learning often.
9. I will strongly recommend others to use STR-texts for learning.

Each of the thirty students in the experimental group turned in a valid answer sheet. Responses to the questionnaire items were scored using a five-point Likert scale, anchored by the end-points "strongly disagree" (1) and "strongly agree" (5). Cronbach's α was employed to assess the internal consistency of the survey. The values were higher than 0.90, suggesting high survey reliability.

We also gave one-on-one semi-structured interviews to ten randomly selected students to explore their interest in using STR-texts in the future as well as to acquire their observations about the potential effectiveness of STR-text for learning. Students were asked the following questions: 1) Did you use STR-text during lectures and for what reasons? 2) Was STR-text useful for learning during lectures and why? Each interview took approximately 30 min; all of the interviews were audio-recorded with the permission of the interviewees and then fully transcribed for analysis. The text segments that met the criteria for providing the best research information were highlighted and coded. Next, the codes were sorted into categories; codes with similar meanings were aggregated together. Established categories produced a framework to illustrate findings relevant to the research questions.

We used a 2 $\times$ 2 statistical analysis (STR intervention by EFL level) to examine if there were interactions between two specific factors, such as STR intervention (i.e. learning with or without STR) and EFL level (i.e. low EFL ability or high EFL ability) on post-test scores and cognitive load. To this end, students in the control and experimental groups were divided into low EFL ability and high EFL ability sub-groups based on their EFL ability: (a) the high EFL ability control group – HACG; (b) the low EFL ability control group – LACG; (c) the high EFL ability experimental group – HAEG; and (d) the low EFL ability experimental group - LAEG. The high EFL ability group included the fifteen top-ranked participants, and the low EFL ability group included the remaining fifteen participants. Differences in learning outcomes and cognitive load between the control and experimental groups were examined by employing an independent sample *t*-test, and differences in the learning outcomes and cognitive load in the low and high ability students in the control and experimental groups were tested using the

Mann-Whitney, non-parametric test. This study set a priori alpha-level (i.e. level of significance) at 0.05 since an alpha level of less than 0.05 is accepted in most educational research as statistically significant.

4. Results

4.1. The effectiveness of STR-texts on learning performance and cognitive load

The means and standard deviations of the assessments and cognitive load are summarized in Table 3. The effects of treatment on post-tests scores and cognitive load are presented in Table 4. The results demonstrate that in Lecture 1 the effect of intervention is significant on learning outcomes, $F(1, 56) = 11.183$, $p < 0.05$, partial eta-squared = 0.166, cognitive load, $F(1, 56) = 9.517$, $p < 0.05$, partial eta-squared = 0.145, and mental effort, $F(1, 56) = 6.195$, $p < 0.05$, partial eta-squared = 0.100. The results show that in Lecture 2 the effect of intervention is also significant on learning outcomes, $F(1, 56) = 31.818$, $p < 0.05$, partial eta-squared = 0.166, cognitive load, $F(1, 56) = 9.308$, $p < 0.05$, partial eta-squared = 0.143, and mental effort, $F(1, 56) = 9.164$, $p < 0.05$, partial eta-squared = 0.141. Our results suggest that students in the experimental group had higher scores on the post-tests and lower cognitive load and mental effort during the two lectures.

A subsequent analysis was carried out using the *t*-test. The results show that there was no significant difference in the EFL ability levels between the control group (M = 772.50, SD = 106.00) and the experimental group (M = 763.17, SD = 139.76), $t = 0.291$, $p > 0.05$. However, the experimental group (M = 13.37, SD = 3.27) outperformed the control group (M = 10.87, SD = 2.98) on the post-test after Lecture 1, $t = -3.093$, $p < 0.05$. The results also show that the experimental group (M = 10.47, SD = 3.48) outperformed the control group (M = 6.80, SD = 1.71) on the post-test after Lecture 2, $t = -5.178$, $p < 0.05$). According to the results, students in the control group had a higher cognitive load in Lecture 1 (M = 2.03, SD = 0.72) than students in the experimental group (M = 1.67, SD = 0.48), $t = 2.325$, $p < 0.05$. In addition, in Lecture 2, the cognitive load of students in the control group (M = 2.30, SD = 0.75) was higher than that of the experimental group (M = 1.93, SD = 0.36), $t = 2.408$, $p < 0.05$. The results also show that students in the control group (M = 2.10, SD = 0.61) expended more mental effort in Lecture 1 than the students in the experimental group (M = 1.77, SD = 0.57), $t = 2.195$, $p < 0.05$. Similar results were found for Lecture 2, in which the mental effort of students in the control group (M = 2.50, SD = 0.82) was higher than that of the experimental group (M = 2.10, SD = 0.40), $t = 2.398$, $p < 0.05$. The results of the *t*-test validate the results of a 2×2 statistical analysis.

4.2. The effectiveness of STR-texts on learning performance and cognitive load for high and low EFL ability students

Descriptive statistics related to learning outcomes and the cognitive load of the low and high EFL ability students in the control and experimental groups are presented in Table 3. The effects of EFL ability on post-tests scores and cognitive load are presented in Table 4. According to the results, in Lecture 1, EFL level has an effect on learning outcomes, $F(1, 56) = 11.787$, $p < 0.05$, partial eta-squared = 0.174. In Lecture 2, EFL level was found to have an effect on learning outcomes, $F(1, 56) = 10.771$, $p < 0.05$, partial eta-squared = 0.161, cognitive load, $F(1, 56) = 9.308$, $p < 0.05$, partial eta-squared = 0.143, and mental effort, $F(1, 56) = 12.473$, $p < 0.05$, partial eta-squared = 0.182.

Our results show that in Lecture 1, STR intervention and EFL level have a significant effect on cognitive load, $F(1, 56) = 6.371$, $p < 0.05$, partial eta-squared = 0.102 and mental effort, $F(1, 56) = 8.920$, $p < 0.05$, partial eta-squared = 0.137. In Lecture 2, STR intervention and EFL level also have a significant effect on cognitive load, $F(1, 56) = 9.308$, $p < 0.05$, partial eta-squared = 0.143 and mental effort, $F(1, 56) = 16.291$, $p < 0.05$, partial eta-squared = 0.225 for the second lecture.

We carried out a subsequent analysis using the non-parametric test. There was no significant difference between LACG (M = 689.33, SD = 85.81) and LAEG (M = 651.67, SD = 112.12), $U = 89.000$, $Z = -0.979$, $p > 0.05$. However, the results demonstrated that LAEG (M = 12.07, SD = 2.99) outperformed LACG (M = 9.73, SD = 3.19) on the post-test after the first lecture, $U = 65.000$, $Z = -1.987$, $p < 0.05$. Furthermore, LAEG (M = 8.93, SD = 2.55) had significantly better performance on

Table 3
The results of the assessments and cognitive load.

	Control group						Experimental group					
	All students		Low ability		High ability		All students		Low ability		High ability	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
EFL ability	772.50	106.00	689.33	85.81	855.67	33.00	763.17	139.76	651.67	112.12	874.67	35.33
Lecture 1												
Post-test	10.87	2.98	9.53	3.07	12.20	2.27	13.37	3.27	12.13	3.00	14.60	3.16
Cognitive load	2.03	0.49	2.27	0.46	1.80	0.41	1.67	0.48	1.60	0.51	1.73	0.46
Mental effort	2.10	0.55	2.40	0.51	1.80	0.41	1.77	0.57	1.67	0.62	1.87	0.52
Lecture 2												
Post-test	6.80	1.71	6.20	1.82	7.40	1.40	10.47	3.48	8.93	2.55	12.00	3.68
Cognitive load	2.30	0.65	2.67	0.62	1.93	0.46	1.93	0.36	1.93	0.46	1.93	0.26
Mental effort	2.50	0.78	3.00	0.76	2.00	0.38	2.10	0.40	2.07	0.46	2.13	0.35

Table 4
The effects of treatment and EFL ability on post-tests scores and cognitive load.

Source	Sum of squares	df	Mean square	F	Sig.	Partial eta-squared
STR intervention						
Lecture 1						
Post-test	93.750	1	93.750	11.183	0.001	0.166
Cognitive load	2.017	1	2.017	9.517	0.003	0.145
Mental effort	1.667	1	1.667	6.195	0.016	0.100
Lecture 2						
Post-test	201.667	1	201.667	31.818	0.000	0.362
Cognitive load	2.017	1	2.017	9.308	0.003	0.143
Mental effort	2.400	1	2.400	9.164	0.004	0.141
EFL level						
Lecture 1						
Post-test	98.817	1	98.817	11.787	0.001	0.174
Cognitive load	0.417	1	0.417	1.966	0.166	0.034
Mental effort	0.600	1	0.600	2.230	0.141	0.038
Lecture 2						
Post-test	68.267	1	68.267	10.771	0.002	0.161
Cognitive load	2.017	1	2.017	9.308	0.003	0.143
Mental effort	3.267	1	3.267	12.473	0.001	0.182
STR intervention * EFL level						
Lecture 1						
Post-test	0.150	1	0.150	0.018	0.894	0.000
Cognitive load	1.350	1	1.350	6.371	0.014	0.102
Mental effort	2.400	1	2.400	8.920	0.004	0.137
Lecture 2						
Post-test	13.067	1	13.067	2.062	0.157	0.036
Cognitive load	2.017	1	2.017	9.308	0.003	0.143
Mental effort	4.267	1	4.267	16.291	0.000	0.225
Error						
Lecture 1						
Post-test	469.467	56	8.383			
Cognitive load	11.867	56	0.212			
Mental effort	15.067	56	0.269			
Lecture 2						
Post-test	354.933	56	6.338			
Cognitive load	12.133	56	0.217			
Mental effort	14.667	56	0.262			

the post-test after the second lecture than LACG ($M = 6.20$, $SD = 1.82$), $U = 43.000$, $Z = -2.908$, $p < 0.05$). According to the results, LACG had higher cognitive load and mental effort during the first ($U = 43.000$, $Z = -2.908$, $p < 0.05$ and $U = 49.500$, $Z = -3.137$, $p < 0.05$ respectively) and second ($U = 46.000$, $Z = -3.169$, $p < 0.05$ and $U = 39.000$, $Z = -3.367$, $p < 0.05$ respectively) lectures.

No significant difference was found between HACG ($M = 855.67$, $SD = 33.00$) and HAEG ($M = 874.67$, $SD = 35.33$), $U = 82.500$, $Z = -1.258$, $p > 0.05$). However, the results showed that HAEG ($M = 14.53$, $SD = 3.25$) outperformed HACG ($M = 12.33$, $SD = 2.38$) on the post-test after the first lecture, $U = 62.000$, $Z = -2.109$, $p < 0.05$). HAEG ($M = 12.00$, $SD = 3.68$) also performed significantly better on the post-test after the second lecture than HACG ($M = 7.40$, $SD = 1.40$), $U = 29.000$, $Z = -3.488$, $p < 0.05$).

When we compared LACG with HACG, results showed that in Lecture 1, LACG exhibited lower learning outcomes ($U = 60.000$, $Z = -2.198$, $p < 0.05$) but higher cognitive load ($U = 66.000$, $Z = -2.609$, $p < 0.05$) and mental effort ($U = 54.000$, $Z = -3.013$, $p < 0.05$) compared to HACG. In additions, in Lecture 2, a nearly significant difference ($U = 69.000$, $Z = -1.840$, $p = 0.066$) was found in learning outcomes, and a significant difference in cognitive load ($U = 46.000$, $Z = -3.169$, $p < 0.05$) and mental effort ($U = 33.500$, $Z = -3.669$, $p < 0.05$) was found between LACG and HACG. That is, LACG had lower learning outcomes and higher cognitive load and mental effort.

Regarding LAEG and HAEG, a significant difference was revealed only with respect to learning outcomes. That is, HAEG outperformed LAEG on the post-tests after both Lecture 1 ($U = 61.500$, $Z = -2.124$, $p < 0.05$) and Lecture 2 ($U = 56.500$, $Z = -2.332$, $p < 0.05$).

4.3. Student perceptions regarding the usefulness of STR-texts for learning

The questionnaire results (see Table 5) show that most students in the experimental group had high perceptions toward the usefulness of STR-texts for learning during both the first ($M = 4.26$, $SD = 0.57$) and second ($M = 4.33$, $SD = 0.48$) lectures. They also stated high behavioral intentions toward use of STR-texts for learning in the future in the questionnaires after Lecture 1 ($M = 4.02$, $SD = 0.62$) and Lecture 2 ($M = 4.13$, $SD = 0.64$).

Table 5
Results of the questionnaire survey.

Dimension	Lecture 1		Lecture 2	
	M	SD	M	SD
Usefulness of STR-texts for learning	4.26	0.57	4.33	0.48
Behavioral intentions to use STR-texts	4.02	0.62	4.13	0.64

Table 6
Categories to use STR-texts for learning.

#	Category	Description	Low ability		High ability	
			%	Lecture	%	Lecture
1	For following the instructor	Students read STR-texts while listening to the instructor to build connection between what they hear and read.	93%	Lecture 1 and 2	–	–
2	For enhancing understanding	Students, who lack listening skills, or students, who had better reading skills than listening, read STR-texts to understand the lecture content better.	80%	Lecture 1 and 2	–	–
3	For confirming what the lecturer has said	When students could not understand the meaning of some sentences that they hear, they referred to STR-texts.	33%	Lecture 1 and 2	20% 67%	Lecture 1 Lecture 2
4	For clarifying some words	Students referred to STR-texts when the instructor pronounced some words unclearly, she spoke not loud enough or too fast.	27%	Lecture 1 and 2	13% 53%	Lecture 1 Lecture 2
5	For making up missed information	Students read STR-texts when they were distracted and missed some important information.	20%	Lecture 1 and 2	–	–

The results of the interviews with the students in the experimental group revealed how useful STR-texts can be for learning during lectures with content given at different levels of difficulty. Categories, their descriptions, what and when students used STR-texts are presented in Table 6. According to the table, students used STR-texts to follow the instructor (93% of LAEG during both lectures), to enhance understanding (80% of LAEG during both lectures), to confirm what the lecturer said (33% of LAEG during both lectures, 20% of HAEG during the first lecture, and 67% of HAEG during the second lecture), to clarify some words (27% of LAEG during both lectures, 13% of HAEG during the first lecture, and 53% of HAEG during the second lecture), and to make up missed information (20% of LAEG during both lectures).

5. Discussion

In this study, our first assumption was that presenting STR-texts during lectures in a foreign language can be useful so that students can learn content better and as a result outperform those who study without STR support. The related studies tested this assumption, but they had some research gaps, such as the validity of empirical results. The results of our study supported our assumptions with sound objective evidence. Our statistical results showed that STR-texts positively affected student learning. Although students in both the control and experimental groups had equal EFL ability, students in the experimental group outperformed those in the control group on the post-test after both lectures. Our second assumption was that lectures in a foreign language result in a heavy cognitive load on the working memory of students due to the complex information processes involved, but this load can be managed if STR-texts are presented during lectures. However, this assumption had not been tested previously in STR-related studies and was supported by the results of our study. The experimental results also showed that STR-texts are useful not only in facilitating and promoting learning but also for keeping students from being cognitively overloaded. We found that students in the control group had higher cognitive load and mental effort during both lectures than students in the experimental group.

The questionnaire results supported the statistical evidence. According to the questionnaire results, in general, students perceived STR-texts as useful for their learning. During the interviews (see interview extracts 1–6 in the Appendix), students confirmed the usefulness of STR-texts for the following reasons: First, most students in LAEG (93%) claimed that while slides provide the key points of the lectures, STR-texts contain all information from lectures and are therefore more useful for learning. These students preferred to read STR-texts to enhance their understanding of the lecture content. Particularly, with STR-texts, students could follow the lectures by listening to the lecturer and reading the STR-texts simultaneously. This behavior enabled them to build connections between what they heard and read. Most students in LAEG (80%) mentioned that STR-texts helped them better understand the lectures and indicated that STR-texts made them feel more confident and less anxious when attending lectures in English. 33% of LAEG after both lectures, 20% of HAEG after the first lecture, and 67% of HAEG after the second lecture mentioned that STR-texts helped them to confirm what the lecturer had said. According to 27% of LAEG during both lectures, 13% of HAEG during the first lecture, and 53% of HAEG during the second lecture, STR-texts were also useful for clarifying key points of the lectures when the instructor pronounced some words unclearly, did not speak loudly enough, or spoke too fast. The students claimed that when they could not understand the meaning of some sentences when they heard them, they would refer to the STR-text for clarification. In addition, students read STR-texts when they were distracted and missed some important information (20% LAEG during both lectures).

When we consider different language abilities, our statistical results suggest that the STR-texts were very important for LAEG so as to improve their learning performance and prevent them from being cognitively overloaded during lectures in a foreign language. The results showed that when LACG learned without STR-texts they had lowest learning performance and highest cognitive load. However, when LAEG learned with STR-texts their performance was better and cognitive load was lower than that of students in LACG. LAEG even had the same low cognitive load and mental effort as HAEG. On the other hand, LACG had higher cognitive load and mental efforts than HACG.

In the interviews (see interview extracts 1–7 in Appendix), LAEG students mentioned that they paid attention to STR-texts during both lectures. STR-texts helped them in the same ways as discussed earlier, i.e. to follow the instructor (93% during both lectures), to enhance understanding (80% during both lectures), to confirm what the lecturer said (33% during both lectures), to clarify some words (27% during both lectures), and to make up missed information (20% during both lectures). Students with better reading skills than listening skills read STR-texts instead of listening to the instructor. Even though the first lecture was not very difficult, and the students' language ability matched its difficulty level, they still relied on STR-texts because they felt more confident and less anxious with STR support. In the second lecture, students mentioned that they seldom paid attention to the slides and focused mostly on the STR-texts when listening to the lecture. This was due to their EFL ability; students in the LAEG mentioned that due to their poor EFL skills, they could not understand some parts of the lecture. However, when the STR-tests were presented, they could understand the lecture content better. Therefore, LAEG students were more dependent on STR-texts and thus, due to the availability of STR-texts, their performance was significantly higher, and their cognitive load and mental effort was lower compared to the LACG students.

With regard to high EFL ability, our results showed that HAEG outperformed HACG on the post-test for both lectures. This finding suggests that STR-texts were useful for students in the HAEG. However, there was no significant difference in cognitive load and mental effort between the two groups. This is due to the students' high EFL ability level. Kalyuga (2014) suggested that the amount of cognitive load depends on how well the learners already know the learning content. All high EFL ability students had some prior knowledge already and could understand the lecture content; this is why all of them had both low cognitive load and mental effort.

Interviews with HAEG students provide additional evidence to support our findings (see interview extracts 8–10 in the Appendix). High EFL ability students in the experimental group admitted that STR-texts were useful during lectures to confirm what the lecturer said and to clarify some words. This behavior helped to enhance comprehension. It should be noted that more students in the HAEG used STR-texts during the second lecture. Students in the HAEG said that the first lecture was very easy to comprehend, so STR-texts were not needed. These students mentioned that listening to the lecturer was enough for them to understand the lecture content. However, when the difficulty level of the lecture was higher (like in the second lecture), the STR-texts suddenly became very useful. This result is supported by the interview results (see Table 6); that is, to confirm what the lecturer said, 20% of high EFL ability students used STR-texts during the first lecture, and 67% of them used them during the second lecture and to clarify some words 13% of high EFL ability students used STR-texts during the first lecture and 53% during the second to lecture content. Based on this finding, it is suggested that STR-texts should be provided to students of different EFL ability levels with some considerations. For example, texts generated during advanced level lectures can be shown to students with low and high EFL ability levels in order to enhance their comprehension. However, in intermediate level lectures, STR-texts need to be provided only to those who feel they will need them. In this way, students who do not need STR-texts will not be overloaded as a result of the learning content being presented in different forms.

Similar results about benefits of STR-text for learning have been reported elsewhere (Hwang et al., 2012; Kuo et al., 2012; Ryba et al., 2006; Shadiev et al., 2016; Wald & Bain, 2008). In addition, our results are in line with related theories. According to the cognitive theory of multimedia learning (Mayer, 2009), when we presented verbal and visual content (i.e. instructor speech and texts generated by STR from the speech input) simultaneously, it contributed to better processing of learning information on the part of the students. The verbal and visual processing systems were involved and coordinated to help process and integrate learning content in English as a foreign language presented in two forms. Students were able to build referential connections between created models and prior knowledge, and both visual and verbal mental models were accessed. This is why presenting STR-texts was very useful, and after presenting learning content in verbal and visual forms simultaneously, students performed better. Cognitive load theory (Plass et al., 2003) also supports our findings. Since students attended lectures in a language that was not their native language, they were cognitively overloaded (i.e. the control students). The reason for this is that such lectures involve high interactivity of elements (Plass et al., 2003), and information processes during lectures are complex (Keysar et al., 2012). Chen and Chang (2009) urged that in such lectures, students have to receive and retain information in working memory and integrate it with what follows, all the while continually adjusting their understanding with prior knowledge. In addition, learning content usually contains many new and unfamiliar concepts (Brunken et al., 2003; Sweller et al., 1998). All of these cause a heavy cognitive demand. This is why we provided students in the experimental group with STR-texts. Students received learning information in verbal and visual forms, which helped diminish information processing demands and prevented them from being cognitively overloaded. Verbal and visual working memories were employed during information processing instead of verbal working memory only, and students cognitive capacity was therefore increased (Clark & Mayer, 2011).

One interesting finding regarding the application of STR technology was derived from the interviews (see interview extract 11 in the Appendix). A few students in the experimental group mentioned that they were not used to multi-media lectures (i.e. including video, slides, and STR-texts). This led their attention to focus on all of the media at the same time at the beginning of the experiment (i.e. during the first lecture). As a result, the students could not comprehend the lecture content efficiently. According to the modality principle proposed by Mayer and Moreno (1998), the same information presented in auditory and written forms makes it redundant and gives rise to a split-attention effect and a higher cognitive load. However, the experimental students in this present study admitted that after watching 1 or 2 lectures, when they got used to the multiple media approach, they only focused on the STR-texts and then comprehended the lecture content better. Based on this finding, it is suggested that instructors need to ensure that students have enough experience to learn with multi-media lectures; i.e. students need to watch at least 1–2 multi-media lectures. Such experiences will allow them to adapt to lectures in which STR-texts are presented and to find the strengths and limitations of using STR-texts for learning.

Our last assumption was that students would be cognitively overloaded during lectures in a foreign language with STR-texts if the lecture content difficulty level matched their language ability due to the redundancy and expertise reversal effects, but if the lecture content was challenging to the students with regard to language, they would not be cognitively overloaded. For example, the first lecture at the intermediate level matched the language ability of the low EFL ability groups, and the second lecture matched the language ability of the high EFL ability groups; therefore, the low EFL ability experimental group (in Lecture 1) and the high EFL ability experimental group (in Lecture 2) would potentially be cognitively overloaded due to exposure to STR-texts. This assumption has not been tested previously in STR-related studies, and it was partially supported by our experimental results.

Our results show that LAEG had lower cognitive load and mental effort compared to LACG during the first and second lectures. That is, no redundancy and expertise reversal effects took place during the first lecture when the same information was presented verbally and visually even when the lecture content difficulty level matched the language ability of students in the LAEG. Cognitive load and mental effort of students in the LAEG were also low during the second lecture compared to LACG. As for high EFL ability students, there was no significant difference in cognitive load and mental effort between students in the control and experimental groups.

Cognitive load theory (Plass et al., 2003) explains these findings. The redundancy effect takes place when the same information is presented to learners in two forms simultaneously, i.e. verbal (i.e. a lecturer's speech) and visual (STR-texts). On the other hand, due to the expertise reversal principle, learning content that is highly effective with novice learners may not be effective and may result in increased cognitive load when used with more knowledgeable learners (Kalyuga, 2014). However, in case of LAEG, learning content presented in two forms was useful for their learning and also did not result in increased cognitive load. Perhaps, this was because the lectures were in a foreign language, which is a special case as mentioned by Clark and Mayer (2011), where redundant visual text becomes acceptable and even necessary. LAEG relied on information presented both in visual and verbal forms to understand it better because they had low language ability and lacked prior knowledge. This finding was supported by the interviews (see interview extracts 2, 5, and 12–13 in Appendix).

On the contrary, HAEG already had prior knowledge, so information in two modalities could become not only redundant but even counterproductive since processing it may have required additional cognitive resources. However, our results showed otherwise. Perhaps, this is because students in HAEG did not use the STR-texts frequently and only in cases where they could not understand the meaning of some sentences or when the instructor pronounced some words unclearly or spoke softly or too fast, so HAEG referred to STR-texts to confirm what the lecturer said and to clarify some words. In addition, only some students reported such learning behavior (please refer to Table 6 for the percentages). This finding was supported by the other evidence obtained from the interviews (see interview extracts 8–10, and 14 in the Appendix).

6. Conclusions

In this study, two important issues that have not been addressed in previous STR-related research were uncovered, and an attempt was made to remedy them. In addition, three assumptions were tested. The following main findings were obtained: First, students who used STR-texts outperformed those who did not use them. Lectures in English caused less cognitive load for low EFL ability students when they used STR-texts. Second, students perceived that STR-texts were useful for learning and they had positive behavioral intentions to use STR-texts for learning in the future. Depending on the students' EFL ability, the STR-texts could be invaluable in helping them to follow the instructor (93% of LAEG during both lectures), to enhance understanding (80% of LAEG during both lectures), and to confirm what the lecturer says (33% of LAEG during both lectures, 20% of HAEG during the first lecture, and 67% of HAEG during the second lecture). Some other reasons to use STR-texts were also demonstrated, e.g. the instructor's accent, pronunciation, speed, or volume could cause comprehension problems (27% of LAEG during both lectures, 13% of HAEG during the first lecture, and 53% of HAEG during the second lecture), or perhaps students were distracted and missed some important information (20% of LAEG during both lectures). Finally, most low EFL ability students (93%) used STR-texts for both the intermediate and advanced level lectures, whereas most high EFL ability students only used STR-texts for the advanced lecture (67%) although some of them did use it for the intermediate level (20%). When the lecture content difficulty matched the experimental students' EFL ability (i.e. lecture 1 and low EFL ability/lecture 2 and high EFL ability), they still performed better compared to the control students. Experimental students whose EFL ability level matched the lecture content difficulty still relied on/paid attention to STR-texts in those lectures.

Based on the results of this study, several suggestions and implications for the teaching and research community in the field can be made. It is suggested that applying STR-texts in lectures that are delivered in English can facilitate students' learning and prevent students from being cognitively overloaded. However, educators and researchers need to consider that STR-texts are useful for low EFL ability students for lectures at both intermediate and advanced levels, while STR-texts should be provided to high EFL ability students only for lectures at the advanced level. The following indicators can be useful for designing or choosing STR-based lectures with appropriate levels of difficulty: (1) a lecture's difficulty level matches to or is higher than language ability of students; (2) listening skills of students are low; (3) reading skills of students are better than their listening skills; (4) lecture contains unfamiliar vocabulary and terminology; (5) students are not confident about attending lectures in a foreign language; (6) the instructor has a strong accent because s/he lectures in a foreign language, and (7) the instructor does not speak loudly enough or speaks too fast. Furthermore, it is suggested that students need to have more experience attending lectures in which STR-texts are presented. In this way, students can find the strengths and limitations of STR-texts in regard to learning and can therefore fully utilize them for learning and to be able to focus to a greater degree on STR-texts from the beginning in order to avoid being cognitively overloaded by information presented in multiple media, thus enabling them to comprehend lecture content more efficiently.

As a result of this study, new knowledge has been added to the field. First, we provided valid empirical evidence that demonstrates the effectiveness of STR-texts on learning performance and cognitive load. Second, we demonstrated how language ability affects the benefits of STR-texts on learning performance and cognitive load during lectures at different difficulty levels. Third, we made several suggestions and implications for the community in the field.

Some limitations of this study need to be acknowledged. The relatively small sample size involved in this study is one limitation. Research with a greater number of participating students is necessary in the future in order to make broader generalizations regarding the results. Short-term exposure to STR-texts is another limitation. In addition, this study only covered two lectures on general domain topics, not on a specific one, e.g. Biology or History. This issue will be addressed in a future study to make findings relevant to "real-world" learning scenarios and settings in which STR-texts are used longer-term.

Future research can extend this work in several ways. First, the effectiveness of STR-texts on learning performance and cognitive load can be explored by collecting and analyzing different sets of data. For example, recording brain wave signals or eye movement during learning may generate data which can be used to measure cognitive load and related mental processes more objectively. Second, future research may focus on investigating how other characteristics of students, such as learning preferences and gender, affect the benefits of STR-texts on learning performance and cognitive load. Third, application of STR technology to support learning and instruction in a foreign language can be broadened. For example, in the future, students may apply STR to generate texts during learning activities, such as group discussion sessions in order to determine how these STR-texts are useful for learning.

Acknowledgement

This research is partially supported by the "International Research-Intensive Center of Excellence Pro-gram" of NTNU and the Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301, MOST 103-2511-S-006-007-MY3, and MOST 103-2511-S-006-002-MY3.

Appendix. Extracts from the interviews with the experimental students

(1) I used STR-texts because I could understand the lecture content better.
(2) I feel nervous and a bit scared when I attend lectures in English because of my language skills. With STR-texts, I do not feel nervous, and I feel confident.
(3) I read STR-texts and listened to the instructor simultaneously to follow the lecture.
(4) The instructor pronounced words unclearly or didn't speak loudly enough sometimes, and in this situation, STR-texts helped me to follow her lectures.
(5) My language skills are not so good, so I read STR-texts during both lectures even though the first lecture was at the intermediate level.
(6) I could review lecture slides and STR-texts in the second lecture, but I rarely paid attention to the slides and read STR-texts all the time.
(7) My reading skills are better than my listening skills, so I like to read instead of listening. I was very happy to attend lectures with STR-texts.

(8) I did not always use STR-texts during the lecture because listening to the instructor was enough for me to understand the content. I read STR-texts when I experienced unfamiliar vocabulary, or I was not able to understand some parts of the lecture.

(9) I was unable to hear the instructor clearly a few times, so I had to refer to STR-texts.
(10) I read STR-texts mostly during the second lecture because it was more difficult.

(11) Because this is my first experience, I was confused at the beginning when I heard the instructor and received STR-texts and slides. It took some time for me to get experienced and benefit from each media type.

(12) I feel it is easy to learn during lectures when STR-texts are shown.
(13) It was not too hard to process lecture content with STR-texts.
(14) My language ability is good enough, so I do not need STR-texts.

References

Bain, K., Basson, S., Faisman, A., & Kanevsky, D. (2005). Accessibility, transcription, and access everywhere. *IBM Systems Journal, 44*(3), 589–603.
Barnes, B. D., & Lock, G. (2010). The attributes of effective lecturers of English as a foreign language as perceived by students in a Korean university. *Australian Journal of Teacher Education, 35*(1), 139–152.
Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61.
Camiciottoli, B. C. (2005). Adjusting a business lecture for an international audience: a case study. *English for Specific Purposes, 24*(2), 183–199.
Chen, I. J., & Chang, C. C. (2009). Cognitive load theory: an Empirical study of anxiety and task performance in language learning. *Electronic Journal of Research in Educational Psychology, 7*(2), 729–746.
Clark, R. C., & Mayer, R. E. (2011). *E-Learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: John Wiley & Sons.
Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative*. Boston, MA: Pearson Education.
Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2014). *How to design and evaluate research in education*. New York: McGraw-Hill.
Huang, Y. M., Huang, Y. M., Liu, C. H., & Tsai, C. C. (2013). Applying social tagging to manage cognitive load in a Web 2.0 self-learning environment. *Interactive Learning Environments, 21*(3), 273–289.
Huang, Y. M., Liu, C. L., Shadiev, R., Shen, M. H., & Hwang, W. Y. (2015). Investigating an application of speech-to-text recognition: a study on visual attention and learning behavior. *Journal of Computer Assisted Learning, 31*(6), 529–545.
Hwang, W. Y., Shadiev, R., Kuo, T. C. T., & Chen, N. S. (2012). Effects of Speech-to-Text Recognition application on learning performance in synchronous cyber classrooms. *Journal of Educational Technology & Society, 15*(1), 367–380.
Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 576–597). New York: Cambridge University Press.
Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The Foreign-language effect thinking in a foreign tongue reduces decision biases. *Psychological Science, 23*(6), 661–668.
Kramer, B. S., Walker, A. E., & Brill, J. M. (2007). The underutilization of information and communication technology-assisted collaborative project-based learning among international educators: a Delphi study. *Educational Technology Research and Development, 55*(5), 527–543.
Kuo, T. C. T., Shadiev, R., Hwang, W. Y., & Chen, N. S. (2012). Effects of applying STR for group learning activities on learning performance in a synchronous cyber classroom. *Computers & Education, 58*(1), 600–608.
Leitch, D. (2008). *Gift atlantic liberated learning high school pilot project: A study of the transfer of speech recognition technology from university classrooms to high school classrooms*. Phase III Report. Nova Scotia, Canada: Saint Mary's University Press.
Mayer, R. E. (2009). *Multimedia learning*. New York: Cambridge University Press.
Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.
Miller, L. (2007). Issues in lecturing in a second language: lecturer's behaviour and students' perceptions. *Studies in Higher Education, 32*(6), 747–760.
Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Psychologist, 38*(1), 1–4.
Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior, 19*(2), 221–243.
Poulsen, R., Hastings, P., & Allbritton, D. (2007). Tutoring bilingual students with an automated Reading Tutor that listens. *Journal of Educational Computing Research, 36*(2), 191–221.
Ranchal, R., Taber-Doughty, T., Guo, Y., Bain, K., Martin, H., Robinson, J., et al. (2013). Using speech recognition for real-time captioning and lecture transcription in the classroom. *IEEE Transactions on Learning Technologies, 6*(4), 299–311.
Ryba, K., McIvor, T., Shakir, M., & Paez, D. (2006). Liberated learning: analysis of university students' perceptions and experiences with continuous automated speech recognition. *Journal of Instructional Science and Technology, 9*(1), 1–19.
Shadiev, R., Hwang, W. Y., Chen, N. S., & Huang, Y. M. (2014). Review of speech-to-text recognition technology for enhancing learning. *Journal of Educational Technology and Society, 17*(4), 65–84 (SSCI).
Shadiev, R., Hwang, W. Y., & Huang, Y. M. (2013). Investigating learning strategies of using texts generated by Speech to Text Recognition technology in traditional classroom. In *Proceedings of the AECT International Conference on the Frontier in e-Learning Research* (pp. 279–286). Taiwan: National Central University & AECT.
Shadiev, R., Hwang, W. Y., Huang, Y. M., & Liu, C. J. (2016). Investigating applications of speech to text recognition for face to face seminar to assist learning of non-native English participants. *Technology, Pedagogy and Education, 25*(1), 119–134.
Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.
Wald, M., & Bain, K. (2008). Universal access to communication and learning: the role of automatic speech recognition. *International Journal Universal Access in the Information Society, 6*(4), 435–447.