Applied Cognitive Psychology, *Appl. Cognit. Psychol.* **28**: 653–660 (2014)
Published online 16 July 2014 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.3050

Multimedia Learning in a Second Language: A Cognitive Load Perspective

RICHARD E. MAYER[1]*, HYUNJEONG LEE[2] and ALANNA PEEBLES[1]

[1]*University of California, Santa Barbara, USA*
[2]*University of Seoul, Seoul, Korea*

Summary: What can be done to help college students who are not native speakers of English learn from computer-based lessons that are presented in English? To help students access the meaning of spoken words in a slow-paced 16-minute narration about wildlife in Antarctica, a representational video was added that showed the scenes and animals being described in the narration (Experiment 1). Adding video resulted in improved performance of non-native English speakers on a comprehension test (d = 0.63), perhaps because the video improved access to word meaning without creating extraneous cognitive load. To help students perceive the spoken words in a fast-paced 9-minute narrated video about chemical reactions, concurrent on-screen captions were added (Experiment 2). Adding on-screen captions did not improve performance by non-native English speakers on comprehension tests, perhaps because learners did not have available capacity to take advantage of the captions. Implications for cognitive load theory are discussed. Copyright © 2014 John Wiley & Sons, Ltd.

Objective and rationale

In college classrooms across the United States, there are students who must learn from lectures that are not presented in their native language. What can be done to help college students who are not native English speakers learn from lessons presented in English? For example, consider the demands on a visiting international student from Korea or China or Japan who is taking a course presented in English at a U.S. university, or even consider the demands on a U.S. college student whose first language is not English.

The goal of the present set of experiments is to test the effectiveness of two instructional methods for helping college students who are not native English speakers learn from computer-based presentations in English—adding redundant video to an audio lecture (Experiment 1) and adding redundant captions to a narrated video lecture (Experiment 2). Adding redundant video represents an attempt to help clarify word meaning (without creating extraneous cognitive load), whereas adding redundant captions is an attempt to help learners parse the incoming auditory stream (but with the possibility of creating extraneous cognitive load). In particular, these experiments examine when redundancy can help or hinder learning for students who receive instruction in their second language.

Although much previous research has focused on learning a second language, including learning a second language with multimedia instruction (Plass & Jones, 2005; Wang, 2011), the present study examines a different issue concerning learning academic content in one's second language (Gu, 2013; Min, 2008). In short, our focus is on learning in a second language rather than second language learning.

Theoretical framework

First, consider students who listen to an audio podcast on Antarctica. Students learning academic content presented in their second language (i.e., non-native speakers) may have weak vocabularies that do not allow them to automatically access word meaning during listening, so they have to expend cognitive resources on trying to infer word meaning from context. This depletes limited cognitive resources that could have been used to build a mental representation of the content and thereby detracts from reading comprehension. In order to support students in accessing word meaning, one technique is to add redundant video that depicts the animals and actions described in the narration, as summarized in the first column of Table 1. Adding video is intended to reduce the need to engage in searching for word meaning (as shown in the second column of Table 1) while not overloading the visual channel (as indicated in the third column of Table 1).

Second, consider students who watch a narrated video on chemical reactions presented via computer. Students learning academic content in their second language who have not automated their phonological processing of English sounds may have difficulty in segmenting the incoming flow of sounds into discreet words and therefore must allocate their limited cognitive processing resources to consciously perceiving each word. This depletes cognitive resources needed for deeper processing and thereby detracts from prose comprehension. In order to support students in perceiving each word, one technique is to add redundant captions to the bottom of the screen, which reproduce each phrase as it is spoken. As summarized in the second column of Table 1, adding on-screen captions is intended to preserve word availability (as spoken words are transitory), making it easier for students to encode the words. However, as shown in the third column of Table 1, adding captions can create extraneous cognitive processing—that is, cognitive processing that does not contribute to the instructional objective and wastes precious processing capacity. In particular, extraneous processing is created when learners must split their visual attention between two places at the same time (i.e., the video and the caption).

Research on redundancy

Research on multimedia learning has yielded a collection of evidence-based principles for the design of multimedia

* Correspondence to: Richard E. Mayer, Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106, USA.
E-mail: mayer@psych.ucsb.edu

Table 1. Cognitive load implications of two aids to multimedia learning in a second language

Instructional aid	Intended to	Load implications
Add representational video to narration	Help access word meaning	Low load
Add on-screen captions to narrated video	Preserve word availability	High load

instructional messages, including the *redundancy principle*, which states that people learn better people from graphics with corresponding spoken words than from graphics with corresponding spoken words and redundant printed words (Mayer, 2009). A recent review by Mayer and Fiorella (2014) found that in 16 out of 16 experimental comparisons, students performed better on a transfer test following a lesson with graphics and spoken words rather than an identical lesson with printed words added to the graphics and spoken words, yielding a median effect size of $d = 0.86$ (Austin, 2009; Craig, Gholson, & Driscoll, 2002; Jamet & Le Bohec, 2007; Kalyuga, Chandler, & Sweller, 1999, 2000; Leahy, Chandler, & Sweller, 2003; Mayer, Heiser, & Lonn, 2001; Moreno & Mayer, 2002a, 2002b; Mousavi, Low, & Sweller, 1995). In another recent review of redundancy effects, Kalyuga and Sweller (2014) came to the same conclusion that adding redundant material to a lesson is detrimental to learning.

The theoretical explanation for the redundancy principle is that adding redundant printed text to narrated graphics creates *extraneous processing*. In the case of redundant on-screen captions, extraneous processing occurs when (i) the learner tries to reconcile two verbal streams to make sure the printed words correspond to the spoken words and (ii) the learner scans between the words in the caption to the corresponding part in the graphic above. Finally, redundant printed text can create split attention in the visual channel in which learners may miss something in a fast-paced video if they spend too much time reading the printed words.

Is there ever a situation in which redundancy can help learning? Mayer and Johnson (2008) reported that students performed better on a retention test (but not transfer) when a narrated animation on lightning formation was supplemented by placing a few key words next to the corresponding part of a graphic. In this case, the printed words minimized extraneous processing because only one or two words appeared, and fostered essential processing by helping the learner select relevant words and connect them with the relevant part of the graphic.

A new area where the redundancy could be helpful involves learning academic content presented in one's second language. On the basis of cognitive load theory (Sweller, Ayres, & Kalyuga, 2011) and the cognitive theory of multimedia learning (Mayer, 2009), the redundancy facilitation hypothesis predicts a reverse redundancy effect in scenarios where redundant material can support basic cognitive processing that is not yet automated in non-native speakers while minimizing extra cognitive load. This kind of scenario (i.e., adding redundant video to audio) is examined in Experiment 1 in which a slow-paced redundant video is added to an audio that contains many unfamiliar animal names. The video serves as a form of sheltered instruction by helping students access word meaning (e.g., showing a penguin in the video when the voice in the narration says 'penguin'), while minimizing any additional cognitive load.

Are there any limits to the redundancy facilitation hypothesis? As has been demonstrated with research on the redundancy effect, adding captions to narrated animation can create demands for extraneous processing that lead to lower scores on transfer tests (Kalyuga & Sweller, 2014; Mayer, 2009; Mayer & Fiorella, 2014). Therefore, this kind of scenario (i.e., adding captions to narrated video) represents a serious challenge when applied as a form of sheltered instruction for students learning in their second language, because students may not have the cognitive capacity to take advantage of the captions. Thus, the redundancy facilitation hypothesis is less likely to apply in this scenario.

Research on learning academic content in one's second language

Sheltered instruction refers to teaching academic content to students in their second language, while providing aids to help them with the mechanics of language processing, but 'valid research is urgently needed' (Hinkel, 2013, p. 311). This study examines two forms of sheltered instruction—adding redundant graphics to an audio lecture and adding redundant captions to a narrated video lecture.

Research on adding redundant graphics shows that adult native speakers generally do not benefit from having redundant illustrations (also called representational illustrations) added to text (Mayer, 2009; Sung & Mayer, 2012). The rationale is that adults learning in their native language do not need the illustrations to help them access the meaning of common words in text, and thus the picture is more of a distraction than an aid. In contrast, there is preliminary evidence that having access to pictures or video that represents words in a narrative that is in one's second language can improve both vocabulary learning and reading comprehension (Chun & Plass, 1996; Gu, 2013; Jones & Plass, 2002; Plass, Chun, Mayer, & Leutner, 1998). The present study seeks to determine whether these effects can be extended to listening to expository text (rather than reading narrative text) and the use of concurrent redundant video (rather than click-on links to see graphics for words in the narrative).

Research on adding redundant captions shows that adult native speakers generally show poorer comprehension when redundant captions are added to multimedia lessons, presumably because they create demands for extraneous processing (Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). In contrast, the use of subtitles has become a common technique in foreign language courses in which films are presented to students new to the language (d'Ydewallel & Van de Poell, 1999). Previous research using films in laboratory settings has found these subtitles result in better content and vocabulary comprehension for non-native speakers (Lavaur & Bairstow, 2011). The present study aims to close the literature gap by testing the effectiveness of subtitles when the goal is learning new academic content instead of new words.

Predictions

On the basis of the analysis summarized in Table 1, we predict that for college students learning in their second language, adding redundant video to audio will result in improvements in comprehension test performance (Experiment 1) because the benefits of helping students access word meaning do cause extraneous cognitive processing demands. In contrast, we predict that when learning in one's second language, adding redundant captions to narrated video may not result in learning improvements (Experiment 2), because of demands for extraneous processing.

EXPERIMENT 1

In Experiment 1, students who were not native English speakers listened to an audio lecture on Antarctica (audio group) or listened to the same audio lecture while viewing a redundant video that depicted the animals and scenes described in the audio (audio + video group). In this case, the cognitive benefit of redundancy is substantial because the images in the video can help learners access word meaning, whereas the cognitive cost is minimal because the visual channel is not otherwise being used.

Method

Participants and design

The participants were 68 university students at the University of California, Santa Barbara (UCSB), who reported not being native speakers of English. Thirteen participants were international students, and 55 were UCSB students who reported that they had initially learned a language other than English. There were 24 men and 44 women. The mean age was 19.25 years ($SD = 1.42$); the mean rating of English listening ability was 3.75 ($SD = 1.05$) with 1 as *low* and 5 as *high*; and the mean rating of knowledge of Antarctica was 2.10 ($SD = 0.90$), with 1 as *low* and 5 as *high*. In a between-subjects design, 34 students served in the audio group, and 34 students served in the audio + video group.

Materials

The instructional materials consisted of an audio version and an audio + video version of a 16-minute lesson on Antarctica, based on a BBC documentary. The Antarctica lesson contained new information that the participants had not learned before the experiment, and the level of the lesson was appropriate for the participants. The audio version consisted of 1122 words spoken in English by a male speaker with a British accent at a slow pace (i.e., approximately 70 words per minute). The lesson described interesting facts about Antarctica, such as the continent of Antarctica is as large as United States and 90% of all of the world's ice is found there; and the lesson described the activities of animals found in Antarctica such as penguins and whales.

The audio + video version contained the same audio as the audio version but also included video of scenes in Antarctica that corresponded to what was being described in the audio script. For example, when the audio described how female emperor penguins transfer the eggs to male penguins, the video showed a mother penguin move the egg from her pouch to a male penguin's pouch. The video was redundant with the text in the sense that it conveyed the same information as the audio message, but the video differed in the sense that it added concrete images representing the words used in the audio.

The paper-based materials consisted of a summary sheet, comprehension test sheets, and participant questionnaire. The summary sheet contained the following instructions at the top of the page: 'Please summarize the main points in the lesson you just heard. You can use bullet points'. The rest of the page was blank except for the instructions 'PLEASE KEEP WORKING UNTIL YOU ARE TOLD TO STOP' at the bottom of the page. A scoring key with 114 idea units was developed, and participants received one point for each idea unit expressed in their summary, regardless of specific wording or spelling, yielding a possible score of 0 to 114.

The comprehension test sheets contained 12 multiple-choice questions based on factual information in the lesson and inferences based on the facts, such as 'In the Antarctic, ice advances at the rate of ___ per day? a. 100 feet, b. 1000 feet, c. 1 mile, d. 2.5 miles' or 'Where do chinstrap penguins lay their eggs? a. water, b. ice, c. rock, d. trees'. Each question was scored as right (one point) or wrong (zero point), so the minimum score is 0 and the maximum score is 12.

The participant questionnaire solicited information concerning the participant's age, gender, year in college, native language, and Test of English as a Foreign Language (TOEFL) reading score and TOEFL listening score (if they were international students). The participant questionnaire asked students to respond to the following items: 'Please rate your knowledge of Antarctica (before the lesson)' on a 5-point scale ranging from *low* (1) to *high* (5); 'Please rate your English listening ability' on a 5-point scale from *low* (1) to *high* (5); 'How difficult was the lesson you just received?' on a 5-point scale from *easy* (1) to *hard* (5); and 'How much effort did you use for the lesson you received?' on a 5-point scale from *low effort* (1) to *high effort* (5).

The apparatus consisted of three iMac computer systems with 21-in. color monitors and Panasonic headphones.

Procedure

Participants were randomly assigned to treatments and tested in groups of up to three per session in a lab environment. Each participant was seated at a desk in an individual cubicle that blocked visual contact with other participants. First, the experimenter passed out the informed consent form, explained the study, and asked participants to read and sign the informed consent form. Then, the experimenter asked the participants to put on their headphones, and the experimenter initiated each participant's 16-minute lesson—either audio or audio + video—based on random assignment. After the lesson, the experimenter passed out the summary test and told participants they would have 5 minutes to write their answer. Then, the experimenter collected the summary test and passed out the comprehension test, which participants completed without a time limit. Next, the comprehension test was collected, and the participant questionnaire was

distributed with no time limit. Upon completion, each participant was debriefed and excused. Throughout the experiment, the rights of the participants were protected, and applicable human subject guidelines were followed. The entire session lasted around 30 minutes.

Results and discussion

Are the groups equivalent on basic characteristics?

A preliminary step is to determine whether the audio group and the audio + video group are equivalent on basic characteristics. On the basis of *t*-tests with $p < .05$, the groups did not differ significantly on mean age, $t(66) = 1.466$, $p = .147$; prior knowledge score, $t(66) = 1.737$, $p = .087$; and English listening skill rating, $t(66) = -0.116$, $p = .908$. On the basis of chi-square tests with $p < .05$, the groups did not differ significantly on the proportion of men and women $X(1) = 0.066$, $p = .7977$, or on the proportion of international students and regular education students, $X(1) = 0.380$, $p = .537$. We conclude that the groups are equivalent on basic characteristics.

In subsequent sections, *t*-tests (with α at .05) are used to compare the groups. However, there was a marginal difference in prior knowledge score favoring the audio + video group ($M = 2.24$, $SD = 0.99$) over the audio group ($M = 1.85$, $SD = 0.82$), so prior knowledge was used as a covariate in analyses of covariance (ANCOVAs) conducted as follow-up analyses. These analyses produced the same conclusions as the *t*-tests.

Does the audio + video group learn better than the audio group?

According to the redundancy facilitation hypothesis, students learning in their second language learn better when words are accompanied by redundant representational graphics. Consistent with this prediction, the video + audio group ($M = 9.47$, $SD = 1.93$) scored significantly higher on the 12-item comprehension test than the audio group ($M = 8.00$, $SD = 2.74$), $t(66) = 2.56$, $p = .013$, $d = 0.63$. However, the groups did not differ significantly on the summary test, $t(66) = 0.23$, $p = .818$, $d = 0.04$; the video + audio group ($M = 2.85$, $SD = 2.16$) remembered about as many idea units as the audio group ($M = 2.97$, $SD = 2.04$), perhaps because the instructions prompted some students to produce generalities rather than specifics. The same significant difference on the comprehension test is obtained, when the groups are compared in an ANCOVA with prior knowledge as a covariate, $F(1, 65) = 5.051$, $p = .028$. Overall, there is evidence that adding redundant, representational video to audio helped students learn in their second language.

In short, adding video to audio created a *multimedia effect* (Butcher, 2014; Mayer, 2009), in which people learn better from words and pictures than from words alone. However, by extending this finding to learning in a second language, our interpretation is that a cause of the multimedia effect could be that the video helped clarify the meaning of unfamiliar words. This interpretation could be tested in future research by examining whether the results disappear for native English speakers or high English language proficient learners, based on a stronger indicator of English language proficiency.

Audio can create a transient information effect in which learners have access to the words only as they are presented (Leahy & Sweller, 2011; Singh, Marcus, & Ayres, 2012). The transient nature of spoken text can be a particular problem for non-native speakers who need time to search for word meaning, so slow moving redundant graphics may help mitigate this problem by keeping some of the described material on the screen.

Does the audio + video group report more effort and less difficulty than the audio group?

According to the redundancy facilitation hypothesis, a secondary prediction is that students learning in their second language should report less difficulty and more effort when words are accompanied by redundant representational graphics. Although the audio + video group reported lower mean rating of difficulty ($M = 2.32$, $SD = 0.84$) than the audio group ($M = 2.73$, $SD = 1.10$), $t(65) = 1.692$, $p = .095$, $d = 0.42$, this trend did not reach statistical significance. The groups reported the same mean rating for effort, $t(66) = 0.000$, $p = 1.000$, $d = 0.000$. Overall, there is not compelling evidence to support the prediction that students will report less difficulty and more effort in learning the material when background video is added to an audio lesson, but more direct and extensive measures of difficulty and effort are needed (Brunken, Seufert, & Paas, 2010).

EXPERIMENT 2

In Experiment 2, students who were not native English speakers viewed a narrated animation involving chemical reactions (no-subtitle group) or viewed the same narrated animation along with concurrent on-screen captions that were identical to the narration (subtitle group). In this case, the cognitive benefit of redundancy may be minimal because the on-screen text serves mainly to help students perceive the words rather than understand their meaning, whereas the cognitive cost is substantial because the visual channel may become overloaded with both captions and video.

Method

Participants and design

The participants were 73 undergraduate international students studying at the UCSB. They were recruited through flyers posted around campus, and they were offered $10 in compensation. There were 37 participants in the no-subtitle group and 36 participants in the subtitle group. There were 24 men and 49 women, and the mean age was 19.90 years ($SD = 2.364$). The majority of the participants' (83.56%) native language was Chinese: 61 participants listed Chinese as their native language, eight listed Korean, and four listed Japanese. The average TOEFL score was 24.98 ($SD = 2.92$) out of 30, and the average participant's self-rating of their English language listening ability was 4.86 ($SD = 1.07$) out of a scale from 1 to 7. The average time the participants had spent in the United States was 14.67 months ($SD = 27.11$).

Materials and apparatus

Paper-based materials included an informed consent form, six post-test question sheets about the video, and a participant questionnaire. The consent form explained that participants' participation was voluntary, their data were anonymous and private, and they could stop their participation at any time during the study without any repercussions.

Each of the six post-test questions was on a separate piece of paper. One was a retention question:

> Explain why you have a fountain of soda when you add Mentos to Diet Coke. Pretend that you are trying to explain this to someone who doesn't know anything about how this process works. Please try to be as detailed as possible.

The other five were transfer questions including redesign, troubleshooting, prediction, and conceptual modeling questions as suggested by Mayer (2009). Examples are as follows: "In this video, you saw a big soda fountain spurt when Adam added a Mentos to Diet Coke. What could be done to make the height of the fountain smaller?" "Your friend hands you a mysterious piece of candy. You drop it into a bottle of soda that's on your kitchen counter, but only a small fountain spurts out of the bottle. Why wasn't there a big fountain?" "You drop a Mentos in water. What do you expect to happen, and why?"

The questions were scored by two independent raters using a rubric consisting of all possible acceptable answers. Exact wording was not necessary for the participant obtain a point for their answer. The total possible idea units was 7 for retention and 26 for transfer. Scoring was highly reliable, $r = .91$, $p < .001$ for retention and $r = .91$, $p < .001$ for transfer.

The participant questionnaire included questions about age, gender, country of origin, length of time spent in the United States, and TOEFL scores. Prior knowledge was assessed with two questions: a scale from 1 to 7 rating 'Your knowledge of chemistry' and a list of items to check if applicable: 'I have seen this episode of MythBusters before', 'I knew about the Diet Coke and Mentos fountain before I watched this video', 'I know what "Au" means on the periodic table', 'I like to learn about science', and 'I have taken a college-level chemistry class'. Additionally, participants were asked to rate their level of English understanding in the written and spoken form on the 1 to 7 scale, with 1 being *very low* and 7 being *very high*.

Computer-based materials included a 9-minute excerpt of the MythBusters's narrated episode 'Diet Coke and Mentos' with subtitles on the bottom of the screen and an excerpt of the same episode without subtitles. Using iMovie as a video editing software tool, we spliced together slices of the episode to focus on one scientific lesson about the explosive chemical reaction between this soda and candy. In total, there were 1308 spoken words presented at a rate of 142.7 words per minute.

The narrated video began with a demonstration of the Mentos and Diet Coke fountain, and it continued with testing separate ingredients of the Diet Coke and the Mentos to examine which factors contributed to the reaction. It ended with a summary of the findings. The subtitled version had one to two lines of text at the bottom of the screen as white text in a semi-transparent black box. These subtitles were displayed at the same time as the lines were being said as exemplified in Figure 1.

The apparatus consisted of five Dell computers with 17-in. screens with Cyber Acoustic headphones.

Figure 1. Screenshot of subtitle version and no-subtitle version of lessons in Experiment 2

Procedure

Each participant was randomly assigned to either the no-subtitle group or the subtitle group prior to entry into the laboratory. Groups of no more than four participants were tested at a time; each person sat at a computer station with booth walls separating their station from the others. Upon entry, they were thanked for coming into the lab, given a short verbal explanation about the experiment and their participant rights, and then given an informed consent form to look over and sign. Participants were then instructed to put on their headphones and begin the video. When they were done, they received the first question and were instructed to keep working and sit quietly until the experimenter called time. This question was collected, and the next question was distributed; this continued for all the post-questions. The first question was the retention question, for which participants had 4.5 minutes to complete. The next five questions were the transfer questions, for which participants had 2.5 minutes to complete. Participants were not allowed to use a dictionary or re-watch the video. At the end of the time for each question, participants were told that they were allowed to finish their last sentence. After the last transfer questions, the experimenter handed them the paper-based untimed questionnaire. Once all the participants of the group finished, a payment form was filled out for each, and a verbal debriefing was made. The duration of the experiment varied between 30 and 50 minutes. These procedures followed all guidelines set forth in the ethical procedures for using human participants.

Results and discussion

Are the groups equivalent on basic characteristics?

The no-subtitle and subtitle groups were compared to examine if they significantly differed on any basic characteristics. Independent samples *t*-test (with $p < .05$) indicated the two groups did not differ significantly on prior knowledge, self-reported level of English listening comprehension, age, and TOEFL listening scores. A chi-square analysis indicated they did not differ in gender or native language distribution. Thus, there is no evidence that the groups differ on any basic characteristics.

Do the groups differ on retention and transfer scores?

The main goal of this study was to determine the possible learning effects for non-native speakers when multimedia lessons are subtitled. An independent samples *t*-test indicated that there was no significant difference between the retention test score of the no-subtitles group ($M = 2.86$, $SD = 1.60$) and the subtitles group ($M = 3.19$, $SD = 1.61$), $t(71) = -0.875$, $p = .357$, $d = 0.20$. Similarly, an independent samples *t*-test indicated that there was no significant difference between the transfer test score of the no-subtitles group ($M = 7.49$, $SD = 2.62$) and the subtitles group ($M = 7.42$, $SD = 2.97$), $t(71) = 0.107$, $p = .915$, $d = -0.01$.

It is possible that subtitles did not overload the visual channel if learners quickly translate them into inner speech. However, having both printed words and graphics on the screen at the same time has the potential to create a split attention effect (Ayres & Sweller, 2014) during initial stages of processing in which learners cannot be looking at the video at the same time they are looking at the subtitle. Similarly, it is possible that learners do not look at the subtitles, but future research using eye tracking would be needed to directly assess this idea.

Do the groups differ on self-reported enjoyment, effort, and difficulty?

The two groups did not differ significantly on self-reported enjoyment of the lesson, $t(71) = 1.21$, $p = .928$ or perceived difficulty of the lesson, $t(71) = 0.493$, $p = .623$. They did, however, differ on the reported effort they put in into understanding the lesson in which subtitles were associated with exerting less effort to understand the lesson ($M = 3.52$, $SD = 1.36$) than no subtitles ($M = 4.27$, $SD = 1.43$), $t(71) = 2.30$, $p = .024$, $d = 0.43$. The reduced effort of the subtitles group did not translate into improved comprehension.

GENERAL DISCUSSION

Empirical contributions

In Experiment 1, adding redundant video to an audio lecture presented in English improved comprehension test performance for students whose first language was not English, whereas in Experiment 2, adding redundant on-screen captions to a narrated video lecture in English did not improve comprehension test performance for students whose first language was not English.

Theoretical contributions

In Experiment 1, redundancy (in the form of a video that mainly depicted the animals and scenes described in the narration) was intended to provide substantial benefits presumably by helping students access word meaning but was intended to not impose substantial costs because the added material was in the visual channel, which was not being used and therefore was not in danger of becoming overloaded. Importantly, Experiment 1 created a multimedia scenario in which adding video to narration was an advantage. A complementary interpretation is that the results of Experiment 1 represent an extension of the multimedia effect for students learning in their second language.

In Experiment 2, redundancy (in the form of an on-screen caption that repeated the words in a narrated video) was intended to help students who are learning in their second language be able to perceptually discriminate the words; however, according to the redundancy principle (based on native speakers), added subtitles had the potential to compete with the fast-moving video for processing capacity in the visual channel if the learner's cognitive capacity was overloaded. Importantly, Experiment 2 created a redundant scenario in which adding subtitles was neither an advantage nor a disadvantage for students learning in their second language, although the null results do allow for definitive interpretations. An alternative explanation is that the students did not attend to the on-screen text, but this idea could best be tested using eye-tracking methodology.

Practical contributions

Not all forms of sheltered instruction are effective for college students learning in their second language. According to the results of Experiment 1, sheltered instruction can be effective when it involves adding video to narration, even when the video simply portrays the scenes and objects described in words. In contrast, according to the results of Experiment 2, sheltered instruction is not necessarily effective when it involves adding subtitles to a fast-paced narrated video. The results of Experiment 2 do not provide support for the proposal that the solution to learning in a second language is to use subtitles.

Limitations and future directions

Future research should investigate whether the effects obtained in Experiment 1 apply to native English speakers or to students with high English proficiency, based on a strong measure of English language proficiency. For the current study, some participants may have had relatively high levels of English listening skill due to their acceptance into UCSB as international students, or due to life experiences outside of home. Many of the international participants were Chinese and had a degree of exposure to English in their childhood and from their early academic courses. It is possible that different results would be found in Experiment 2 with students who had lower English language skill. Self-report instruments for prior knowledge, language proficiency, mental effort, and cognitive load may have limited validity, so interpretations based on these measures should be made with caution.

Experiment 2 was different from Experiment 1 in terms of content and context, so future research is needed to replicate our findings with other materials. The null findings in Experiment 2 do not allow for causal conclusions, so further research is needed to pinpoint the circumstances under which adding subtitles is helpful or not helpful to students learning in a second language. Although we did not find support for the redundancy facilitation hypothesis in Experiment 2, interpretation of a null effect is facilitated by replication studies that examine the consistency of findings (Shavelson & Towne, 2002). In addition, further research should be done to assess different learning situations, especially in a more ecological setting. This study was not done with material from an existing academic course, and as such, the lesson itself may have not been as conducive to learning as a well-refined lecture may be. Participants may not have been as motivated as students in a classroom to try their best, because there were no serious consequences of their performance, and the post-test was conducted immediately after participants had watched the video.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Office of Naval Research. Experiment 1 was conducted while Hyunjeong Lee was on sabbatical leave at the University of California, Santa Barbara. Rebecca Clark, Jessica Pelayo, and Aaron Wertheimer assisted in data collection and scoring in Experiment 1. Experiment 2 is based on a senior honors thesis by Alanna Peebles, which was supported by an ORCA grant from UCSB College of Letters and Science.

REFERENCES

Austin, K. A. (2009). Multimedia learning: Cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules. *Computers and Education*, *53*, 1339–1354.

Ayres, P., & Sweller, J. (2014). The split attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed; pp. 206–226). New York: Cambridge University Press.

Brunken, R., Seufert, T., & Paas, F. (2010). Measuring cognitive load. In J. L. Plass, R. Moreno, & R. Bruken (Eds.), *Cognitive load theory* (pp. 181–202). New York: Cambridge University Press.

Butcher, K. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed; pp. 174–205). New York: Cambridge University Press.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, *80*, 183–198.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*, 428–434.

Gu, Y. (2013). Second language vocabulary. In J. Hattie, & E. M. Anderson (Eds.), *International guide to student achievement* (pp. 307–309). New York: Routledge.

Hinkel, E. (2013). Language teaching curricula. In J. Hattie, & E. M. Anderson (Eds.), *International guide to student achievement* (pp. 310–311). New York: Routledge.

Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology*, *32*, 588–598.

Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, *86*, 546–561.

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edn; pp. 247–262). New York: Cambridge University Press.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351–371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126–136.

Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, *46*(6), 455–462.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation, and the transient information effect. *Applied Cognitive Psychology*, *25*, 943–951.

Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology*, *17*, 401–418.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous cognitive processing in multimedia learning: Coherence, signaling, redundancy, spatial continuity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edn; pp. 279–315). New York: Cambridge University Press.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*, 380–386.

Mayer, R. E., Heiser, H., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*, 187–198.

Min, C. (2008). The problems and solutions of English immersion education. *Journal of English Education*, *7*(1), 109–123.

Moreno, R., & Mayer, R. E. (2002a). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, *94*, 156–163.

Moreno, R., & Mayer, R. E. (2002b). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*, 598–610.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*, 319–334.

Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 467–488). New York: Cambridge University Press.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, *90*, 25–36.

Shavelson, R. J., & Towne, L. T. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Singh, A., Marcus, N., & Ayres, P. (2012). The transient information effect: Investigating the impact of segmentation on spoken and written text. *Applied Cognitive Psychology*, *26*, 848–853.

Sung, E., & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*, *28*, 1618–1625.

Sweller, J., Ayres, P., & Kalyuga, S., (2011). *Cognitive load theory*. New York: Springer.

Wang, M. (2011). Learning a second language. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 127–147). New York: Routledge.

d'Ydewallel, G., & Van de Poell, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, *28*(3), 227–244.