



HAL
open science

Reducing the split-attention effect of subtitles during video learning: might the use of occasional keywords be an effective solution?

Salomé Cojean, Nicolas Martin

► To cite this version:

Salomé Cojean, Nicolas Martin. Reducing the split-attention effect of subtitles during video learning: might the use of occasional keywords be an effective solution?. *L'Année psychologique*, 2021, Vol. 121 (4), pp.417-442. 10.3917/anpsy1.214.0417 . hal-03616023

HAL Id: hal-03616023

<https://hal.science/hal-03616023>

Submitted on 22 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title:

Reducing the split-attention effect of subtitles during video learning: might the use of occasional keywords be an effective solution?

Réduire l'effet d'attention divisée des sous-titres pendant l'apprentissage en vidéo : les mots-clés ponctuels sont-ils une solution efficace ?

Running title:

Effects of subtitles and keywords during video learning

Effets des sous-titres et des mots-clés en apprentissage vidéo

Salomé Cojean & Nicolas Martin

Corresponding author:

Salomé Cojean

Affiliation: Laboratoire de Recherche sur les Apprentissages en Contexte (LaRAC, EA 602),
Univ. Grenoble Alpes, France

Address:

Salomé Cojean

1251 Avenue Centrale,
38400 Saint-Martin-d'Hères,
France

Email: salome.cojean@univ-grenoble-alpes.fr

Nicolas Martin

Affiliation: LIG, CNRS, Grenoble INP, Univ. Grenoble Alpes, France

Address:

Nicolas Martin

700 Avenue Centrale,
38400 Saint-Martin-d'Hères,
France

Email: nicolas.martin6@univ-grenoble-alpes.fr

Reducing the split-attention effect of subtitles during video learning: might the use of occasional keywords be an effective solution?

Abstract

Learning via videos presents many positive aspects (e.g., animation, multi-modality) but also has some constraints. For example, when subtitles are provided, a split-attention effect could occur between the oral narration, written text, and visual illustration. The presentation of only a few written keywords instead of subtitles may be a good solution in terms of how to guide learners into their information selection process. In the current study, 96 participants were distributed among four experimental conditions. They were shown a 12-minutes video with or without subtitles, and with or without highlighted information (i.e., keywords). The results showed no effect of subtitles, but keywords had a negative impact on content memorization, comprehension, and on the time allocated to learning. The results are discussed in terms of metacognition and learners' strategies. It is then hypothesized that learners did not use keywords as relevant scaffolds. Instead of being guided into the selection process, learners may have considered that the keywords replaced it, and overestimated their learning.

Keywords: Video, subtitles, keywords, split-attention, learning

Réduire l'effet d'attention divisée des sous-titres pendant l'apprentissage en vidéo : les mots-clés ponctuels sont-ils une solution efficace ?

Abstract

Les vidéos pédagogiques présentent des bénéfices (e.g., animations, multi-modalité) mais également des contraintes à prendre en compte. Par exemple, en présence de sous-titres,

un effet d'attention divisée peut apparaître entre l'audio, le texte écrit et les illustrations. Une solution serait de présenter des mots-clés à la place des sous-titres afin d'éviter une redondance inutile et de guider l'apprenant à sélectionner l'information pertinente. Dans cette étude, 96 participants ont appris une vidéo de 12 minutes avec ou sans sous-titres, et avec ou sans information mise en saillance (i.e., mots-clés). Les résultats n'ont montré aucun effet des sous-titres, mais les mots-clés ont un effet négatif sur la mémorisation, la compréhension et le temps d'apprentissage. Ces résultats sont discutés en termes de stratégies mises en place. Les apprenants n'auraient pas utilisé les mots-clés pour les guider dans la sélection de l'information, mais les auraient considérés comme remplaçant ce processus, surestimant ainsi leur apprentissage.

Mots-clés : Vidéo, sous-titres, mots-clés, attention divisée, apprentissage

1. Introduction

The use of pedagogical videos has increased over the last 20 years, a trend marked by three main factors: 1) Massive Online Open Courses (MOOC), generally in the form of video-recorded lectures (Giannakos, 2013; Schacter & Szpunar, 2015); 2) flipped classrooms involving videos for the lesson to be studied before the class (Haagsman et al., 2020; van der Meij & Böckmann, 2021); and 3) distance learning in schools and universities due to the COVID-19 pandemic (Dietrich et al., 2020; Kim, 2020). Presenting the pedagogical content via video may have some benefits, but it also has some constraints compared to traditional text-based content. Research on video learning has grown since 2007 (Giannakos et al., 2014), but questions remain on how to efficiently design the videos, taking into account new technological advances.

Indeed, technical progress in artificial intelligence (LeCun et al., 2015) has allowed new possibilities to improve education. Relying on deep learning methods (Goksel & Bozkurt, 2019), it is possible to automatically computerize some tasks, such as the classification, segmentation, or generation of various types of documents (e.g., text, images). In the context of pedagogical videos, some tasks previously requiring human intervention such as subtitles (e.g., Qiu, 2020; Shi et al., 2019) or summaries (e.g., Yousefi-Azar & Hamey, 2017) can now be automatically generated. Even if such operations are now technically feasible, their contributions to learning must still be evaluated.

Although many studies have been conducted on the effect of subtitles in video in foreign language situations (e.g., d'Ydewalle & De Bruycker, 2007; Guichon & McLornan, 2008; Guillory, 1998; Kanellopoulou, 2019; Lavaur & Bairstow, 2011; Lee et al., 2013; Liao, Kruger, et al., 2020; Szarkowska & Bogucka, 2019; van der Zee et al., 2017; Winke et al., 2013; Zanón, 2006), the processes involved may differ (Kanellopoulou et al., 2019; Lin et al., 2016) from learning with same-language subtitles (or video captions; see Gernsbacher, 2015). In the current

study, the impact of same-language subtitles and keywords on learning will be investigated. This topic seems particularly relevant, as designers might be tempted to insert them mechanically into video-based environments. In the light of recent studies about signaling information in videos, keywords will be associated with markers in the timeline in the present study. This implies less frequent keywords than those commonly studied in the literature about partial redundancy but may maximize their impact on learning.

1.1.Video-based learning

1.1.1. Multimedia learning

Multimedia learning implies that information is delivered by texts (written or heard) and pictures (Mayer, 2014b). According to Paivio's dual coding theory (1990), two separated channels are devoted to information processing: one is a verbal system devoted to language processing (e.g., written or oral words) and the other is an imagery system devoted to non-verbal information (e.g., visual objects, sounds, tactile sensations) in which information can be processed in parallel (see Figure 1). These channels are also considered as capacity-limited. The theory relies on the general assumption of working memory as capacity-limited, and the possibility of overloading it with too much incoming information for it to process. This cognitive overload may be detrimental to learning (Paas & Sweller, 2014; Sweller et al., 2019). In the cognitive load theory, three types of cognitive load are usually distinguished: the intrinsic cognitive load that refers to the complexity of information; the extraneous cognitive load related to information presentation; and the germane cognitive load, defined as the resources left to process information (Paas & Sweller, 2014; Sweller et al., 2019, p. 201). The aim of pedagogical design is thus to: 1) reduce the extraneous cognitive load to 2) increase the germane cognitive load to 3) process the intrinsic cognitive load (Paas & Sweller, 2014).

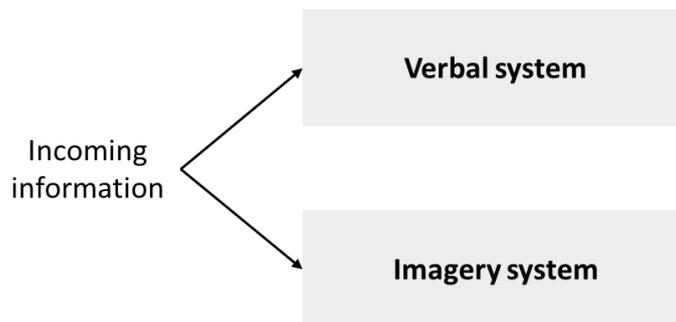


Figure 1. Paivio's distinction between verbal and imagery systems for information processing

Figure 1. Distinction entre le système verbal et le système imagé pour le traitement de l'information selon le modèle de Paivio

According to the Cognitive Theory of Multimedia Learning proposed by Mayer (2014a), and in line with Paivio's work as outlined above, incoming information is processed through two capacity-limited channels. But more precisely, this process implies the active participation of learners, who select the relevant information from the document, organize it into mental representations, and then integrate these representations with their previous knowledge. This succession of steps is common to various other models, and may be designated as the Selection-Organization-Integration model (Fiorella & Mayer, 2016). However, theories about multimedia learning have not been specifically designed for video-based learning. Therefore, some particularities related to the video format need to be considered.

1.1.2. Benefits of the video format to learning

The video format combines two major benefits over traditional written pedagogical documents (van der Meij & van der Meij, 2014) that are not exclusive to videos, but are combined in this format. The first is that videos make it possible to show animated illustrations, which are beneficial compared to static pictures according to the literature (for a meta-analysis,

see Höffler & Leutner, 2007). This makes it possible to represent dynamic content, and gives the learner access to the perception of changes in time and/or space (Lowe & Schnotz, 2014; Ploetzner et al., 2020; van der Meij & van der Meij, 2014).

The second benefit of the video format is that visual and auditory information can be delivered simultaneously (van der Meij & van der Meij, 2014). This benefit is related to the way the information is processed. On the basis of Paivio's work, learning models such as the Cognitive Theory of Multimedia Learning (Mayer, 2014a) have emerged, alongside the Integrated Model of Text and Picture Comprehension (Schnotz, 2014). According to these models, information is still processed in two separated capacity-limited channels, but a preliminary sensory level is added in which the two channels distinguish between auditory and visual modalities. After this, information is processed in verbal and pictorial models in Mayer's model, and in depictive and descriptive subsystems according to Schnotz's model. This distinction at a cognitive level between verbal and non-verbal channels appears very similar to that described by Paivio (1990) for verbal and imagery systems (see Figure 2).

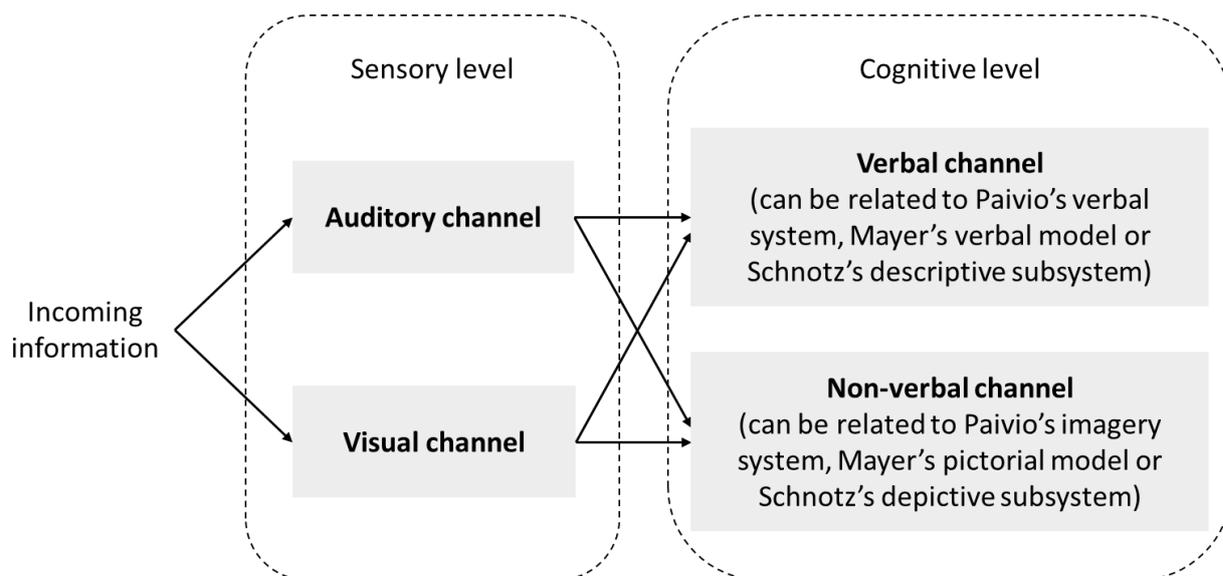


Figure 2. Information processing in two channels at sensory and cognitive levels

Figure 2. Traitement de l'information en deux canaux au niveau sensoriel et au niveau cognitif

Thus, cognitive overload could arise both at cognitive (as conceptualized by Paivio) and sensory levels. Presenting auditory and visual information could increase the effective working memory capacity, by not overloading either of the channels (Low & Sweller, 2014; Sweller et al., 2019). In the multimedia learning literature, this beneficial effect is called the modality effect (for a meta-analysis, see Ginns, 2005). Furthermore, in videos, auditory information is usually in the form of verbal narration, and visual information is usually provided as animated pictures. So, irrespective of the theoretical distinction considered to exist between the two channels (auditory/visual or verbal/non-verbal), the video format should be particularly adapted to solicit both of them. In summary, the video format delivers information in a particular way that can lead to better information processing. Liao et al. (2020) proposed a model specifically including a subtitles process. This model also presents benefits from the parallel use of visual

and auditory channels. However, the video format also brings some constraints, that must now be considered.

1.1.3. Negative aspects of the video format to learning

Along with the positive aspects of video over written documents, some negative impacts of videos must also be noted. The main concern related to information delivered transiently (Höffler & Leutner, 2007; Leahy & Sweller, 2011; Merkt et al., 2011). Contrary to static information (e.g., written text), when individuals watch some animated content, they need to integrate the delivered information with previous information that is no longer available (Hasler et al., 2007; Schwan & Riempp, 2004; Singh et al., 2012). This kind of information processing is cognitively costly (i.e., it increases the extraneous cognitive load; see Wong et al., 2019), and could exceed individuals' resources (Bétrancourt et al., 2003; Hasler et al., 2007; Spanjers et al., 2012; Sweller et al., 2019). Indeed, the process of maintaining access to representations and integrating new information into these representations mainly relies on the working memory capacity, which is limited (Wiley et al., 2014).

Another constraint of the video format is related to the difficulty of retrieving specific information. Rather than looking for information in a written document by visually browsing it, individuals need to navigate through the timeline of the video until they find the information they seek (Zhang et al., 2006). However, it appears that the ability to identify content's organization in a document is a key component in information processing (Sanchez et al., 2001). This ability may also rely on working memory capacity: indeed, individuals with higher working memory capacity may better maintain and integrate information during retrieval activity (Banas & Sanchez, 2012).

Overall, the aims when designing a video-based learning environment remain the same as those of multimedia learning to 1) avoid cognitive overload; and 2) facilitate the processes

of selection, organization (both promoting information seeking process), and integration (constrained by transience and limited working memory capacity) of relevant information. For this purpose, the presence of subtitles needs to be analyzed in view of these aims. It appears that, in line with the previously presented models, subtitles may induce cognitive overload. The next section will explain how full subtitles may be more detrimental than beneficial to the learner, but partial subtitles such as keywords may resolve some of the negative aspects.

1.2. Presence of subtitles in a video-based environment

1.2.1. Subtitles and split-attention

Contrary to expectations based on the modality effect, in a study conducted by Leahy and Sweller (2011), participants had a lower level of learning performance when visual illustration was associated with oral narration rather than written text. According to the authors, this result can be explained by the transitory nature of oral narration, that makes it too difficult to process in working memory when the speech is too long. In this specific situation, the negative transience effect would hinder the beneficial modality effect. A shorter, less complex oral narration or written text would be easier to process. Thus, it might appear that adding written text (i.e., subtitles) to oral narration could limit the negative transitory aspect of information, but the combination of written text and oral narration (with an illustration) could actually be more detrimental than beneficial. Indeed, a negative split-attention effect could then occur when learners have to simultaneously mentally integrate different sources of information (Hasler et al., 2007). A subtitled video would provide three concurrent sources of information at the same time: oral narration, written text, and visual illustration (Zanón, 2006). According to Paivio's model (1990), this implies a risk of cognitive overload in the verbal system, while in Mayer's (2014a) and Schnotz's (2014) models the risk of overload is situated in the visual channel (van der Zee et al., 2017; Yue et al., 2013). This situation may lead to losing the benefit

of the modality effect (i.e., the use of two channels instead of one to avoid overload). Furthermore, the association of oral narration and written subtitles may lead to redundancy, since these two sources deliver identical information (Adesope & Nesbit, 2012; Sorden, 2005). Generally, redundancy is considered as negative, individuals learn better when the pedagogical document is constructed of pictures and spoken words rather than pictures, spoken and written text presented simultaneously (Jamet & Le Bohec, 2007; Kalyuga et al., 2004; Mayer & Fiorella, 2014).

Lastly, the presence of subtitles has specific constraints related to written content: learning is mainly driven by the text. In a study led by Hannus and Hyönä (1999, experience 2), individuals were given a document with static illustration and written text. The authors' analysis of eye movements showed that learners spent only 6% of learning time inspecting the illustrations. Overall, when written text is presented, individuals start reading before alternating between text and pictures, and ultimately spend more time on the text (Schmidt-Weigand et al., 2010). However, in video, visual information is transient (Ayres et al., 2019). Even if it can be considered less transient than oral speech, adding subtitles implies adding partially transient written information, and the reading rhythm is no longer controlled by the learner (Kruger et al., 2015). Attention being mainly focused on the subtitles may lead to not enough time being devoted to inspecting illustrated information that may quickly disappear.

So, the main constraints of subtitles in videos are that a split-attention effect can occur between the text and pictures, and that the text is considered as the principal source of information. The literature generally agrees on the potential negative effect of redundancy that full subtitles may provide with oral narration, even if some studies contradict these results (e.g., Gernsbacher, 2015; Szarkowska et al., 2016). A way to limit the amount of time allocated to redundant information could be to present only keywords associated with the oral narration in writing (Adesope & Nesbit, 2012).

1.2.2. Full versus partial subtitles

In studies by Mayer and Johnson (2008), learners were exposed to a diagram accompanied by oral narration. Written keywords were proposed or not. The presence of redundant keywords was found to lead to better memorization of the information. In a meta-analysis conducted by Adesope and Nesbit (2012), the degree of redundancy between oral narration and written text seemed to moderate the redundancy effect. Indeed, partial redundancy (e.g., oral narration associated with written keywords only) may lead to better learning than full redundancy (e.g., oral narration, and identical written text). The positive effect of keywords may be explained by the fact that keywords are supposed to guide learners in selecting relevant information (Adesope & Nesbit, 2012; Mayer & Johnson, 2008; McCrudden et al., 2014). As it corresponds to the first step of the learning process according to the Selection-Organization-Integration model (Fiorella & Mayer, 2016; Mayer, 2014a), facilitating information selection may be particularly useful in improving learning (de Koning et al., 2009).

In fact, this highlighting benefit attributed to keywords could be related to the definition of signaling devices in written documents (e.g., titles, typographical variations) according to Lorch (1989), aiding the identification of important information. When dealing with transient content (i.e., video), it may be particularly helpful for learners to know which information is important but also when it is presented, which could decrease extraneous cognitive load (de Koning et al., 2009). Further, studies on the effect of titles in videos (when associated with markers in the timeline, e.g., Cojean & Jamet, 2017, 2018) have shown that they may facilitate navigation in the video and the retrieval of specific information (which is one of the major constraints of videos). Thus, it can be supposed that occasional keywords with markers in the timeline could be beneficial in the same way in dealing with transient information and facilitating the retrieval of information during learning, especially during reviewing activity.

1.3. The current study

The current study aims to investigate the impact of full subtitles and keywords (i.e., partial subtitles) on learning in video. When focused on the redundant effect, previous studies appeared either to focus on the effect of subtitles (e.g., Jamet & Le Bohec, 2007) or on the effect of keywords (e.g., Mayer & Johnson, 2008). In a study conducted by de Koning, van Hooijdonk, and Lagerwerf (2017), the effects of oral speech, written subtitles and on-screen labels on procedural learning were analyzed. However, the animation only lasted 84 seconds. According to results from Leahy and Sweller (2011), difficulties due to transience may depend on the length of the video. Perhaps the video in the study of de Koning, van Hooijdonk, and Lagerwerf (2017) was too short to induce a negative impact of redundancy and split-attention. It may therefore be interesting to analyze the impact of subtitles and keywords when learning using a longer video that might exceed learners' cognitive resources. The current study aims to evaluate these effects through the use of a 12-minute video. Moreover, in the current study, and contrary to de Koning, van Hooijdonk, and Lagerwerf (2017), keywords were associated with markers in the timeline to optimize their beneficial guiding effect. Following a 2x2 experimental design, learners were shown a video with full subtitles, only some keywords, both (subtitles and keywords), or none. Our main hypotheses are that: 1) subtitles may be totally redundant with oral narration, split attention between too many sources of information, and then have a negative impact on learning performance; and 2) occasional keywords may provide only partial redundancy and guide learners into relevant strategies (i.e., the selection of relevant information, navigation) that have an overall positive effect on learning.

In this context, learning corresponds to the increase in knowledge in the long-term memory (Hasler et al., 2007), and it implies the memorization as well as comprehension of the presented information (Eitel et al., 2013). On the one hand, memorization can be defined as the ability to reproduce or recognize presented material (Mayer, 2014b), and is usually measured

through retention tests (Bétrancourt et al., 2003). On the other hand, comprehension corresponds to the interaction between incoming information and prior knowledge (Woodall et al., 1983), and is usually measured through transfer tests (Bétrancourt et al., 2003) in which learners can reuse what they have learned in new contexts.

A series of predictions was made in the present study, as expressed in the following hypotheses:

Hypothesis 1. Effect of subtitles on video learning. Due to the split-attention effect between three concurrent sources of information (oral narration, written text, and visual illustration), full redundancy between two sources (oral narration and written text) and attention being focused on the written text at the expense of visual illustration, the presence of subtitles may be detrimental to the learner's comprehension (Hypothesis 1a) and memorization (Hypothesis 1b) of the video content.

Hypothesis 2. Effects of keywords on video learning. Keywords may help learners to select important information and facilitate the retrieval of specific information. So, the presence of keywords may be beneficial to comprehension (Hypothesis 2a) and the memorization of video content. However, previous studies have shown that the memorization of non-highlighted information could be poorer in the presence of keywords, because it would be considered as less important information (Lorch, 1989; Lorch & Lorch, 1996). Thus, a distinction is made by the learner between important information (i.e., highlighted by keywords when they are provided) and less important information (i.e., not highlighted by keywords). The presence of keywords may be beneficial to the memorization of important information (Hypothesis 2b), but they may have a negative impact on the memorization of less important information (Hypothesis 2c).

2. Method

2.1. Participants

113 students attending the University of Angers (France) participated in the study. Data from 17 participants were excluded from the analyses due to video issues. Of the 96 remaining students (87 women, 9 men), the average age was 20.07 years ($SD = 2.15$). All the participants volunteered to take part in the study, and were recruited via an email sent to all students at the university. The experiment was conducted in accordance with the principles of the Declaration of Helsinki (World Medical Association, 2013).

2.2. Materials and experimental design

Each participant had access to a computer and headphones. A video-based environment was designed to display a video in French about astronomy (TEDx Talks, 2013). The video-based environment is presented in Figure 3, and consisted of the video (lasting 11 minutes and 51 seconds), a table of contents presenting the five main parts of the video, and the timeline with segmentation corresponding to the five chapters. A timer was designed just below the table of contents to indicate how much time was left for learning (maximum 20 minutes). A button reading "Go to the questionnaire" was also displayed below the timer. A table of contents representing the chapters of the video was available in each experimental condition, as were segments in the timeline corresponding to these chapters. All the participants could navigate through the video using the timeline and click on a chapter in the table of contents to directly access it. Participants were evenly and randomly distributed among the four experimental conditions. In the "Subtitles" condition ($n = 24$), subtitles transcribing the full oral narration (i.e., verbatim; Szarkowska et al., 2011) were included below the screen (see Figure 3). In the "Keywords" condition ($n = 25$), seven keywords were available instead of full subtitles. They appeared when they were pronounced by the oral narration and corresponded to the actual written version of the spoken text. No other written information was included except the

keywords. Depending on the keyword, they were displayed on screen for between two and four seconds. Keywords were chosen as they were considered as representing important information (i.e., the key ideas) in the delivered content. According to de Koning et al. (2009), highlighting information in transient content implies that the learner should guide their attention onto which information is highlighted, and when. Furthermore, previous studies in the literature (e.g., Cojean & Jamet, 2017, 2018) have shown that titles only have a beneficial impact on information retrieval and navigation in a video when they are associated with markers in the timeline. Then, markers were situated in the timeline to represent the localization of keywords (see Figure 3). In the "Subtitles and Keywords" condition (n = 23), full subtitles were displayed, and keywords were highlighted in these subtitles. Corresponding markers in the timeline were also available (see Figure 3). Finally, in the "Control" condition (n = 24), the video was displayed without any subtitles or keywords (see Figure 3).

Table des matières

- 1: Introduction
- 2: La vie d'une étoile
- 3: La voie lactée
- 4: Plus loin que notre galaxie
- 5: Plus loin que la dernière galaxie

Il reste 19 min 55 sec

Lorsque vous avez fini votre apprentissage, cliquez sur le bouton ci-dessous :

[Passer au questionnaire](#)

1



02:59/11:51

1 2 3 4 5

TEDuParis, Christophe Gaffard (2013). 400.000 galaxies, et toi et moi, au centre de l'univers.

Table des matières

- 1: Introduction
- 2: La vie d'une étoile
- 3: La voie lactée
- 4: Plus loin que notre galaxie
- 5: Plus loin que la dernière galaxie

Il reste 19 min 55 sec

Lorsque vous avez fini votre apprentissage, cliquez sur le bouton ci-dessous :

[Passer au questionnaire](#)

2



02:59/11:51

1 2 3 4 5

TEDuParis, Christophe Gaffard (2013). 400.000 galaxies, et toi et moi, au centre de l'univers.

Table des matières

- 1: Introduction
- 2: La vie d'une étoile
- 3: La voie lactée
- 4: Plus loin que notre galaxie
- 5: Plus loin que la dernière galaxie

Il reste 19 min 55 sec

Lorsque vous avez fini votre apprentissage, cliquez sur le bouton ci-dessous :

[Passer au questionnaire](#)

3



02:59/11:51

1 2 3 4 5

TEDuParis, Christophe Gaffard (2013). 400.000 galaxies, et toi et moi, au centre de l'univers.

Table des matières

- 1: Introduction
- 2: La vie d'une étoile
- 3: La voie lactée
- 4: Plus loin que notre galaxie
- 5: Plus loin que la dernière galaxie

Il reste 19 min 55 sec

Lorsque vous avez fini votre apprentissage, cliquez sur le bouton ci-dessous :

[Passer au questionnaire](#)

4



02:59/11:51

1 2 3 4 5

TEDuParis, Christophe Gaffard (2013). 400.000 galaxies, et toi et moi, au centre de l'univers.

Figure 3. Screenshots of video-based environment in Control (1), Subtitles (2), Keywords (3) and Subtitles and Keywords (4) conditions

Figure 3. Captures d'écran de l'environnement vidéo pour les conditions Contrôle (1), Sous-titres (2), Mots-clés (3) et Sous-titres et mots-clés (4)

2.3. Measures

2.3.1. Control variables

Interest in the topic and perceived competence. A pre-task questionnaire was proposed to participants before they undertook the learning task. They were asked to indicate from 0 to 10 (on an 11-point Likert scale) their degree of interest in the topic of the video (i.e., galaxies in the universe). They also had to indicate from 0 to 10 (11-point Likert scale) how much they felt competent on this subject. The aim of these two questions was to ensure that participants were equitably distributed among the four experimental conditions.

2.3.2. Learning time

The video lasted 11 minutes and 51 seconds, but learners had 20 minutes to learn it. Then, time was left for the learners to revise the information. They could navigate through the video as much as they wanted during these 20 minutes. In the current study, this could be considered as a measure of learning strategies. If they considered that their learning was done before the end of the 20 minutes, they could go to the post-task questionnaire whenever they wanted. The learning time was measured from when the learners launched the video (i.e., when the timer started) to when they skipped to the post-task questionnaire (i.e., when they clicked on the "Go to the questionnaire" button or directly accessed the questionnaire at the end of the 20 minutes).

2.3.3. Comprehension

A post-task questionnaire was proposed to the learners after the learning task. Five comprehension questions (i.e., transfer questions) were asked (four open-ended questions and one multiple-choice question). The answers were not explicitly given in the video; rather, learners had to infer responses (e.g., “What process in space can be related to the famous Lavoisier sentence “Nothing is lost, nothing is created, everything is transformed”?”). Each question was worth one point. In total, the comprehension questions were scored as five points, then transformed into a percentage. Questions were scored by one rater.

2.3.4. Memorization

In the post-task questionnaire, 13 memorization questions (12 open-ended questions and one multiple-choice question) were asked, for a total of 15 points. Eleven questions were worth one point each, and the other two questions were worth two points each because the expected responses were composed of two parts. All the information to answer these questions was explicitly contained in the video. Six memorization questions (on six points) were directly about important information that the keywords highlighted when they were available. For example, in the “Keywords” and “Subtitles and Keywords” conditions, one of the keywords (or key phrases) was “13 million light years”, because it was considered as important information in the video (see Figure 3). Then, one question was “What is the maximum distance, with a telescope, at which we can see galaxies?”. Seven other questions (on nine points) were about information considered as less important and thus not related to potential keywords. For example, one of those questions was “In its beginnings, what was the temperature of the universe?”, and this information was not related to any keywords. Since scores on important and less important information were not on the same scale, they were transformed into

percentages. As for comprehension questions, memorization questions were scored by one rater.

2.4. Procedure

One to eight participants could perform the experiment at the same time. Once they were installed, each one in front of a computer, the experimenter presented the study and the instructions. Each participant started when they wanted. As was mentioned above, first, the participants answered a pre-task questionnaire about their interest and perceived competence in the topic of the video. Then, they had to read the instructions for the learning task. It was stipulated that they would have 20 minutes to learn the maximum information from the video, and that they could navigate as much as they wanted during those 20 minutes. When they wanted to go to the knowledge test, they would have to click on the "Go to the questionnaire" button. If they had not finished in 20 minutes, the video would stop and they would automatically go to the questionnaire. The post-task questionnaire was presented on the computer screen and was composed of 13 memorization and 5 comprehension questions presented in a random order. At the end of the questionnaire, two demographic questions about age and gender were asked.

3. Results

All statistical analyses were performed using the R statistical software (R Core Team, 2019).

3.1. Control variables

An ANalysis Of VAriances (ANOVA) was performed on the prior interest and perceived competence scores (from 0 to 10) to ensure that the participants were equitably distributed among the experimental conditions. The ANOVA showed no differences between conditions (see Table 1 for descriptive statistics), either for interest in the topic, $F(3, 92) = 0.78$, $p = .510$, $\eta^2 = .02$, or perceived competence, $F(3, 92) = 0.88$, $p = .454$, $\eta^2 = .03$.

Table 1. Descriptive statistics for control variables

Tableau 1. Statistiques descriptives pour les variables contrôles

Condition	Prior interest		Perceived competence	
	M	SD	M	SD
Control	6.29	1.30	2.00	1.22
Subtitles	5.83	2.39	2.67	1.95
Keywords	6.00	1.55	2.16	1.34
Subtitles and Keywords	5.39	2.74	2.43	1.56

3.2. Learning

3.2.1. Learning time

Concerning learning time, the ANOVA showed a significant effect of keywords on learning times, $F(1, 92) = 5.88$, $p = .017$, partial $\eta^2 = .06$. Participants with keywords seemed to spend less time learning the video than participants without keywords (see Table 2 for descriptive statistics). Since the video lasted 11 minutes and 51 seconds (i.e., 711 seconds), participants with keywords spent an average of 2 minutes and 3 seconds (i.e., 123.41 seconds) engaged in reviewing activity, whereas participants without keywords spent 3 minutes and 31 seconds on average (i.e., 211.07 seconds) reviewing information. No effects of subtitles, $F(1, 92) = 0.06$, $p = .806$, partial $\eta^2 = .00$, nor of interaction between keywords and subtitles, $F(1, 92) = 0.38$, $p = .538$, partial $\eta^2 = .00$, were found.

Table 2. Descriptive statistics for learning times

Tableau 2. Statistiques descriptives pour les temps d'apprentissage

Condition	Learning time (in s.)	
	M	SD
Control	929.71	197.49
Subtitles	914.42	220.53
Keywords	819.60	126.46
Subtitles and Keywords	849.22	153.31

3.2.2. Comprehension

The ANOVA showed a significant effect of keywords on comprehension scores, $F(1, 92) = 5.05, p = .027$, partial $\eta^2 = .05$. Participants with keywords seemed to have a lower comprehension score than participants without keywords (see Table 3 for descriptive statistics).

No effects of subtitles, $F(1, 92) = 1.15, p = .287$, partial $\eta^2 = .01$, nor of interaction between keywords and subtitles, $F(1, 92) = 0.18, p = .675$, partial $\eta^2 = .00$, were found.

Table 3. Descriptive statistics for comprehension scores

Tableau 3. Statistiques descriptives pour les scores de compréhension

Condition	Comprehension (%)	
	M	SD
Control	35.42	15.32
Subtitles	41.25	23.83
Keywords	28.00	20.87
Subtitles and Keywords	30.43	18.02

3.2.3. Memorization

Concerning the global memorization score, the results showed a significant effect of keywords, $F(1, 92) = 6.67, p = .011$, partial $\eta^2 = .07$, but no effect of subtitles, $F(1, 92) = 2.50$,

$p = .117$, partial $\eta^2 = .03$, nor any interaction between keywords and subtitles, $F(1, 92) = 0.38$, $p = .538$, partial $\eta^2 = .00$.

Questions about important information (i.e., information highlighted in conditions with keywords) and questions about less important information (i.e., that which was not highlighted by keywords in conditions with keywords) were also distinguished. For questions about important information, the ANOVA showed a significant effect of keywords on memorization scores, $F(1, 94) = 6.64$, $p = .012$, $\eta^2 = .07$. For questions about less important information, the results also showed a significant effect of keywords, $F(1, 94) = 4.63$, $p = .034$, $\eta^2 = .05$.

Overall, participants with keywords seemed to have lower memorization scores than participants without keywords, both for highlighted information and non-highlighted information (see Table 4 for descriptive statistics).

Table 4. Descriptive statistics for memorization scores

Tableau 4. Statistiques descriptives pour les scores de mémorisation

Condition	Memorization (%)		Memorization of important information (%)		Memorization of less important information (%)	
	M	SD	M	SD	M	SD
Control	54.51	17.02	62.50	25.30	49.19	17.80
Subtitles	58.33	20.02	66.32	26.86	53.01	18.61
Keywords	41.13	21.37	46.33	27.86	37.67	20.33
Subtitles and Keywords	50.14	23.26	52.90	33.77	48.31	18.04

4. Discussion

The aim of this study was to test the effects of subtitles and keywords in a video on learning performance. Full subtitles were supposed to have a negative impact on learning due to 1) the split-attention effect between illustration, oral narration, and written text;

2) redundancy between oral and written narration; and 3) learners' attention being focused on the written text instead of illustrations. In contrast, using occasional keywords with markers on the timeline was supposed to be positive since it could 1) highlight relevant information and guide learners to be more efficient in their selection process (Mayer & Johnson, 2008); and 2) facilitate navigation through the video (Cojean & Jamet, 2017).

The results showed no effect of subtitles on learning performance. Hypotheses 1a and 1b are therefore not validated. If a split-attention effect occurred, it had no impact on learning in this study. Two explanative hypotheses could be considered and explored in future works. The first of these is that since learning seems to be generally driven by the text (e.g., Hannus & Hyönä, 1999; Schmidt-Weigand et al., 2010), learners may have focused on subtitles when they were present and totally inhibited information from the oral speech. However, this is not in line with the theoretical concept of redundancy (Mayer & Fiorella, 2014). Secondly, a reverse possibility is that the learners ignored the written text (i.e., subtitles). Maybe it is easier to focus only on oral speech and visual illustrations when the video is long and when subtitles are present at a high frequency, contrary to keywords being less frequent and way more salient in drawing learners' attention. Studies could therefore be conducted with slightly different information delivered via written text and oral speech in order to explore which source was favored in this context. Further, the effect of keywords surprisingly seemed to be negative on learning performance. Learners with keywords had lower comprehension scores than learners without keywords. Hypothesis 2a is thus not validated. Moreover, the presence of keywords had a negative impact on the memorization of less important information as well as on the memorization of important information, thus confirming Hypothesis 2c but not validating Hypothesis 2b. Regarding learning time, it appears that learners with keywords spent less time on the video, which indicates less time reviewing the information that could have been relevant.

An explicative hypothesis could be that learners don't perceive keywords as particularly relevant (de Koning et al., 2009). Also, in the current study, the presence of occasional and salient keywords could have led to the implementation of non-relevant strategies by learners. Keywords are supposed to help the selection of important information by signaling it, but learners may have considered that instead of supporting the selection process, the keywords replaced it. Then, they may overestimate their learning performance, which is consistent with less time spent watching the video. This difficulty for learners in pursuing a relevant strategy (i.e., using keywords to facilitate the selection process) must be explored, and may be explained in terms of self-regulation difficulties (Fernandez & Jamet, 2017; Winne & Hadwin, 2013). Lastly, keywords were supposed to reduce the split-attention effect between the spoken and written text. However, in relation to the results of the current study, it can be hypothesized that keywords caused more split-attention than subtitles due to their low frequency of appearances and disappearances (i.e., only seven instances). In fact, the number and frequency of keywords may be an important factor in the process: a lower number of more salient keywords would be worse than partial subtitles throughout the video. In studies conducted by Mayer and Johnson (2008), the term "keywords" actually referred to partial subtitles present at each slide (i.e., at a high frequency), and were found to have a positive impact on learning. Further, in a study led by Angerbauer, Adel and Vu (2019), full subtitles were compared to compressed subtitles (the latter with about 50% of the words presented in the full subtitles). No differences were found in comprehension between the conditions. The keywords presented in the current study represented far less than 50% of the subtitles. The small number of keywords may have had a negative impact on learners' attention. It may therefore be necessary to find the right ratio between too much redundancy and too little salient information. Further work could also help to establish which information to highlight, as this may guide learners' attention and strategies.

The current study has some limits. First, all the participants were college students. A positive side of this is that students are daily engaged in learning activities, and are probably more likely to deal with video formats than other people. However, to generalize the observed effects, a larger sample of participants would be needed. Further, only one video was used in the proposed learning activity. Replication of the study using different video-based environments and different types of content would be necessary to generalize the conclusions. Next, only one rater scored the memorization and comprehension questionnaires. Even if the rater followed a detailed scale, double blind scoring should be considered in future studies. Also in future studies, additional measures could be recorded as verbalizations to gain access to learners' strategies and self-regulation processes (Fernandez & Jamet, 2017; Greene et al., 2015; Kruger & Steyn, 2014). Eye-tracking measures could also be used to gain insights into learners' visual strategies (e.g., Kruger & Steyn, 2014): it may be the case that a split-attention effect between picture, oral narration, and written text doesn't appear because learners strongly rely on one source rather than another. Additional measures could also explore the attention specifically given to oral narration when written text is read. Finally, the linguistic characteristics of keywords (e.g., length, level of abstraction) could be considered, as well as subtitles and keyword editing (i.e., verbatim or edited text; Szarkowska et al., 2011, 2016).

5. Conclusion

In conclusion, contrary to theoretical redundancy and split-attention effects, the current study observed no impact on learning performance when subtitles are totally redundant with oral narration. However, the presence of some keywords had a negative impact on learning. Learners' attention was probably even more divided between the oral speech, visual illustration, and keywords, because they were only occasional. There is also the possibility that learners were misled by the keywords and did not pursue efficient learning strategies. Future studies should focus on a deeper comprehension of learning strategies and the processes involved in

the presence of keywords. It is also necessary to look for a specific design promoting the potentially beneficial effect of keywords on the selection process.

6. Acknowledgments

The authors would like to thank students and members of the LPPL (Laboratoire de Psychologie des Pays de la Loire) for their material support.

7. Declaration of Conflicting Interest

The authors declare that there is no conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

8. References

- Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology, 104*(1), 250–263.
<https://doi.org/10.1037/a0026147>
- Angerbauer, K., Adel, H., & Vu, N. T. (2019). Automatic Compression of Subtitles with Neural Networks and its Effect on User Experience. *Interspeech 2019*, 594–598.
<https://doi.org/10.21437/Interspeech.2019-1750>
- Ayres, P., Castro-Alonso, J. C., Wong, M., Marcus, N., & Paas, F. (2019). Factors that impact on the effectiveness of instructional animations. In S. Tindall-Ford, S. Agostinho, & J. Sweller (Eds.), *Advances in Cognitive Load Theory* (1st ed., pp. 180–193). Routledge.
<https://doi.org/10.4324/9780429283895-15>
- Banas, S., & Sanchez, C. A. (2012). Working Memory Capacity and Learning Underlying Conceptual Relationships Across Multiple Documents: Working Memory and Underlying Relationships. *Applied Cognitive Psychology, 26*(4), 594–600.
<https://doi.org/10.1002/acp.2834>
- Bétrancourt, M., Dillenbourg, P., & Montarnal, C. (2003). Computer technologies in powerful learning environments: The case of using animated and interactive graphics for teaching

- financial concepts. In E. De Corte, L. Verschaffel, N. Entwistle, & J. J. G. Van Merriënboer (Eds.), *Unravelling basic components and dimensions of powerful learning environments* (pp. 143–157). Elsevier.
- Cojean, S., & Jamet, E. (2017). Facilitating information-seeking activity in instructional videos: The combined effects of micro-and macroscaffolding. *Computers in Human Behavior*, *74*, 294–302. <https://doi.org/10.1016/j.chb.2017.04.052>
- Cojean, S., & Jamet, E. (2018). The role of scaffolding in improving information seeking in videos. *Journal of Computer Assisted Learning*, *34*(6), 960–969. <https://doi.org/10.1111/jcal.12303>
- d’Ydewalle, G., & De Bruycker, W. (2007). Eye Movements of Children and Adults While Reading Television Subtitles. *European Psychologist*, *12*(3), 196–205. <https://doi.org/10.1027/1016-9040.12.3.196>
- de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, *21*(2), 113–140. <https://doi.org/10.1007/s10648-009-9098-7>
- de Koning, B. B., van Hooijdonk, C. M. J., & Lagerwerf, L. (2017). Verbal redundancy in a procedural animation: On-screen labels improve retention but not behavioral performance. *Computers & Education*, *107*, 45–53. <https://doi.org/10.1016/j.compedu.2016.12.013>
- Dietrich, N., Kentheswaran, K., Ahmadi, A., Teychené, J., Bessière, Y., Alfenore, S., Laborie, S., Bastoul, D., Loubière, K., Guigui, C., Sperandio, M., Barna, L., Paul, E., Cabassud, C., Liné, A., & Hébrard, G. (2020). Attempts, successes, and failures of distance learning in the time of COVID-19. *Journal of Chemical Education*, *97*(9), 2448–2457. <https://doi.org/10.1021/acs.jchemed.0c00717>

- Eitel, A., Scheiter, K., & Schüler, A. (2013). How inspecting a picture affects processing of text in multimedia learning. *Applied Cognitive Psychology*, 27(4), 451–461. <https://doi.org/10.1002/acp.2922>
- Fernandez, J., & Jamet, E. (2017). Extending the testing effect to self-regulated learning. *Metacognition and Learning*, 12(2), 131–156. <https://doi.org/10.1007/s11409-016-9163-9>
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, 28(4), 717–741. <https://doi.org/10.1007/s10648-015-9348-9>
- Gernsbacher, M. A. (2015). Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 195–202. <https://doi.org/10.1177/2372732215602130>
- Giannakos, M. N. (2013). Exploring the video-based learning research: A review of the literature. *British Journal of Educational Technology*, 44(6), E191–E195. <https://doi.org/10.1111/bjet.12070>
- Giannakos, M. N., Jaccheri, L., & Krogstie, J. (2014). Looking at MOOCs rapid growth through the lens of video-based learning research. *International Journal of Emerging Technologies in Learning*, 9(1), 35–38. <https://doi.org/10.3991/ijet.v9i1.3349>
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15(4), 313–331. <https://doi.org/10.1016/j.learninstruc.2005.07.001>
- Goksel, N., & Bozkurt, A. (2019). Artificial intelligence in education: Current insights and future perspectives. In *Handbook of Research on Learning in the Age of Transhumanism* (pp. 224–236). IGI Global.
- Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C., & McVea, M. (2015). Domain-specificity of self-regulated learning processing in science and history.

- Contemporary Educational Psychology*, 42, 111–128.
<https://doi.org/10.1016/j.cedpsych.2015.06.001>
- Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*, 36(1), 85–93.
<https://doi.org/10.1016/j.system.2007.11.005>
- Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal*, 89–108.
- Haagsman, M. E., Scager, K., Boonstra, J., & Koster, M. C. (2020). Pop-up questions within educational videos: Effects on students' learning. *Journal of Science Education and Technology*, 29(6), 713–724. <https://doi.org/10.1007/s10956-020-09847-3>
- Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low-and high-ability children. *Contemporary Educational Psychology*, 24(2), 95–123. <https://doi.org/10.1006/ceps.1998.0987>
- Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, 21(6), 713–729.
<https://doi.org/10.1002/acp.1345>
- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17(6), 722–738.
<https://doi.org/10.1016/j.learninstruc.2007.09.013>
- Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology*, 32(4), 588–598.
<https://doi.org/10.1016/j.cedpsych.2006.07.001>
- Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors*, 46(3), 567–581.
<https://doi.org/10.1518/hfes.46.3.567.50405>

- Kanellopoulou, C. (2019). Film Subtitles as a Successful Vocabulary Learning Tool. *Open Journal of Modern Linguistics*, 09(02), 145–152. <https://doi.org/10.4236/ojml.2019.92014>
- Kanellopoulou, Kermanidis, & Giannakoulou. (2019). The Dual-Coding and Multimedia Learning Theories: Film Subtitles as a Vocabulary Teaching Tool. *Education Sciences*, 9(3), 210. <https://doi.org/10.3390/educsci9030210>
- Kim, J. (2020). Learning and teaching online during Covid-19: Experiences of student teachers in an early childhood education practicum. *International Journal of Early Childhood*, 52(2), 145–158. <https://doi.org/10.1007/s13158-020-00272-6>
- Kruger, J.-L., & Steyn, F. (2014). Subtitles and Eye Tracking: Reading and Performance. *Reading Research Quarterly*, 49(1), 105–120. <https://doi.org/10.1002/rrq.59>
- Kruger, J.-L., Szarkowska, A., & Krejtz, I. (2015). Subtitles on the moving image: An overview of eye tracking studies. *Refractory: A Journal of Entertainment Media*, 25, 1–14.
- 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. <https://doi.org/10.1080/00207594.2011.565343>
- Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, 25(6), 943–951. <https://doi.org/10.1002/acp.1787>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lee, M., Roskos, B., & Ewoldsen, D. R. (2013). The Impact of Subtitles on Comprehension of Narrative Film. *Media Psychology*, 16(4), 412–440. <https://doi.org/10.1080/15213269.2013.826119>

- Liao, S., Kruger, J. L., & Doherty, S. (2020). The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension. *Journal of Specialised Translation, 33*, 70–98.
- Liao, S., Yu, L., Reichle, E. D., & Kruger, J.-L. (2020). Using Eye Movements to Study the Reading of Subtitles in Video. *Scientific Studies of Reading, 1*–19. <https://doi.org/10.1080/10888438.2020.1823986>
- Lin, J. J., Lee, Y.-H., Wang, D.-Y., & Lin, S. S. (2016). Reading subtitles and taking notes while learning scientific materials in a multimedia environment: Cognitive load perspectives on EFL students. *Journal of Educational Technology & Society, 19*(4), 47–58.
- Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1*(3), 209–234. <https://doi.org/10.1007/BF01320135>
- Lorch, R. F., & Lorch, E. P. (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology, 88*(1), 38–48. <https://doi.org/10.1037/0022-0663.88.1.38>
- Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 227–246). Cambridge University Press.
- Lowe, R. K., & Schnotz, W. (2014). Animation principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 513–546). Cambridge University Press.
- Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>

- Mayer, R. E. (2014b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 1–24). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. *The Cambridge Handbook of Multimedia Learning*, 279–315.
- Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, 100(2), 380–386. <https://doi.org/10.1037/0022-0663.100.2.380>
- McCrudden, M. T., Hushman, C. J., & Marley, S. C. (2014). Exploring the Boundary Conditions of the Redundancy Principle. *The Journal of Experimental Education*, 82(4), 537–554. <https://doi.org/10.1080/00220973.2013.813368>
- Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2011.03.004>
- Paas, F. G., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 27–42). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.004>
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press.
- Ploetzner, R., Berney, S., & Bétrancourt, M. (2020). A review of learning demands in instructional animations: The educational effectiveness of animations unfolds if the features of change need to be learned. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12476>

- Qiu, X. (2020). Study on Automatic Generation of Teaching Video Subtitles Based on Cloud Computing. In J.-S. Pan, J. Li, P.-W. Tsai, & L. C. Jain (Eds.), *Advances in Intelligent Information Hiding and Multimedia Signal Processing* (Vol. 156, pp. 309–314). Springer Singapore. https://doi.org/10.1007/978-981-13-9714-1_34
- R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sanchez, R. P., Lorch, E. P., & Lorch, R. F. (2001). Effects of headings on text processing strategies. *Contemporary Educational Psychology*, 26(3), 418–428. <https://doi.org/10.1006/ceps.2000.1056>
- Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, 1(1), 60. <https://doi.org/10.1037/stl0000011>
- Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction*, 20(2), 100–110. <https://doi.org/10.1016/j.learninstruc.2009.02.011>
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge University Press.
- Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, 14(3), 293–305. <https://doi.org/10.1016/j.learninstruc.2004.06.005>
- Shi, B., Ji, L., Liang, Y., Duan, N., Chen, P., Niu, Z., & Zhou, M. (2019). Dense Procedure Captioning in Narrated Instructional Videos. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 6382–6391. <https://doi.org/10.18653/v1/P19-1641>

- Singh, A.-M., Marcus, N., & Ayres, P. (2012). The Transient Information Effect: Investigating the Impact of Segmentation on Spoken and Written text: The transient information effect. *Applied Cognitive Psychology*, 26(6), 848–853. <https://doi.org/10.1002/acp.2885>
- Sorden, S. D. (2005). A cognitive approach to instructional design for multimedia learning. *Informing Science: International Journal of an Emerging Transdiscipline*, 8, 263–279.
- Spanjers, I. A. E., van Gog, T., Wouters, P., & Van Merriënboer, J. J. G. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education*, 59(2), 274–280. <https://doi.org/10.1016/j.compedu.2011.12.024>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Szarkowska, A., & Bogucka, L. (2019). Six-second rule revisited: An eye-tracking study on the impact of speech rate and language proficiency on subtitle reading. *Translation, Cognition & Behavior*, 2(1), 101–124. <https://doi.org/10.1075/tcb.00022.sza>
- Szarkowska, A., Krejtz, I., Klyszejko, Z., & Wieczorek, A. (2011). Verbatim, standard, or edited? Reading patterns of different captioning styles among deaf, hard of hearing, and hearing viewers. *American Annals of the Deaf*, 156(4), 363–378.
- Szarkowska, A., Krejtz, I., Pilipczuk, O., Dutka, Ł., & Kruger, J.-L. (2016). The effects of text editing and subtitle presentation rate on the comprehension and reading patterns of interlingual and intralingual subtitles among deaf, hard of hearing and hearing viewers. *Across Languages and Cultures*, 17(2), 183–204. <https://doi.org/10.1556/084.2016.17.2.3>

- TEDx Talks. (2013, November 28). *400,000 galaxies, et toi et moi, au centre de l'univers: Christophe Galfard at TEDxParis.*
<https://www.youtube.com/watch?v=OsmIEnPORSa>
- van der Meij, H., & Böckmann, L. (2021). Effects of embedded questions in recorded lectures. *Journal of Computing in Higher Education*, 33(1), 235–254.
<https://doi.org/10.1007/s12528-020-09263-x>
- van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & Education*, 78, 150–159.
- van der Zee, T., Admiraal, W., Paas, F. G., Saab, N., & Giesbers, B. (2017). Effects of subtitles, complexity, and language proficiency on learning from online education videos. *Journal of Media Psychology*, 29(1), 18–30. <https://doi.org/10.1027/1864-1105/a000208>
- Wiley, J., Sanchez, C. A., & Jaeger, A. J. (2014). The Individual Differences in Working Memory Capacity Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed.).
- Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, 97(1), 254–275. <https://doi.org/10.1111/j.1540-4781.2012.01432.x>
- Winne, P. H., & Hadwin, A. F. (2013). nStudy: Tracing and Supporting Self-Regulated Learning in the Internet. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (Vol. 28, pp. 293–308). Springer New York.
https://doi.org/10.1007/978-1-4419-5546-3_20
- Wong, M., Castro-Alonso, J. C., Ayres, P., & Paas, F. (2019). The effects of transient information and element interactivity on learning from instructional animations. In S.

- Tindall-Ford, S. Agostinho, & J. Sweller (Eds.), *Advances in cognitive load theory* (1st ed., pp. 80–88). Routledge. <https://doi.org/10.4324/9780429283895-7>
- Woodall, W. G., Davis, D. K., & Sahin, H. (1983). From the boob tube to the black box: Television news comprehension from an information processing perspective. *Journal of Broadcasting*, 27(1), 1–23. <https://doi.org/10.1080/08838158309386469>
- World Medical Association. (2013). Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, 310(20), 2191. <https://doi.org/10.1001/jama.2013.281053>
- Yousefi-Azar, M., & Hamey, L. (2017). Text summarization using unsupervised deep learning. *Expert Systems with Applications*, 68, 93–105. <https://doi.org/10.1016/j.eswa.2016.10.017>
- Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty? *Journal of Educational Psychology*, 105(2), 266–277. <https://doi.org/10.1037/a0031971>
- Zanón, N. T. (2006). Using subtitles to enhance foreign language learning. *Porta Linguarum: Revista Internacional de Didáctica de Las Lenguas Extranjeras*, 6, 4.
- Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, 43(1), 15–27. <https://doi.org/10.1016/j.im.2005.01.004>